

System Analysis Using Simulation for Manufacturing Gauge R&R Study

César Rodríguez

A Thesis

in

The Department

of

Mechanical and Industrial Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science

in

Industrial Engineering

at

Concordia University

Montreal, Quebec, Canada

June, 2015

© César Rodríguez, 2015

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the Thesis prepared

By: César Augusto Rodríguez Gallegos

Entitled: System Analysis Using Simulation for Manufacturing Gauge R&R Study

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Industrial Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. M. Hojjati Chair

Dr. M. Kazemi Zanjani Examiner

Dr. A. Awasthi Examiner

Dr. M. Chen Supervisor

Approved by Dr. S. Narayanswamy
Chair of Department or Graduate Program Director

Amir Asif
Dean of Faculty

Date Jul 09 2015

Abstract

System Analysis Using Simulation for Manufacturing Gauge R&R Study

César Rodríguez

Manufacturing environment relies on fulfillment of product specifications and customer satisfaction. Quality characteristics of products are measured for assessing its accordance to requirements, but measurement error might mask true process performance. Measurement system variability must be small with respect to product specifications as well to process variation. Total variation in a manufacturing process is a combination of part-to-part variation and measurement variation. Prior to endeavor in process analysis and enhancement it is highly recommended to perform a measurement system capability study. The analysis of the measurement system has to take into account three basic components, appraiser, equipment and product. Measurement system analysis (MSA) relies on statistical tools, such as Gauge R&R study, to ensure that the measurement system is in acceptable conditions to monitor the manufacturing process. Main objective of the thesis is to analyze how sensitive are the confidence intervals of variability components and capability criteria to the design of a Gauge R&R study. We conducted extensive numerical experiments using simulation. Findings will be of help as guidance to practitioners when dealing with parameter allocation decisions in a Gauge R&R study.

Acknowledgments

To God, for blessing my life and allowing me finish my Master program satisfactorily.

I would like to express my gratitude to my supervisor, Dr. Mingyuan Chen, for his priceless support and guidance. I sincerely appreciate his dedication and hard work through this two years of research.

Financial assistance received from the government of Ecuador and Concordia University is highly appreciated.

I want to thanks my beloved parents, Fernando and Gladys, for all their unconditional support, love and motivation.

And special thanks to mi abuelito Fernando, who always inspires and encourages me to go ahead.

Contents

Chapter 1 Introduction	1
1.1 Motivation	1
1.2 Objectives of the thesis	2
1.3 Research in the thesis	2
1.4 Thesis structure	3
Chapter 2 Literature Review	4
2.1 Gauge repeatability and reproducibility study	4
2.1.1 Introduction	4
2.1.2 Gauge R&R study and its applications	5
2.1.3 Gauge R&R study design	7
2.1.4 Gauge R&R study methods	8
2.1.5 Confidence intervals for measurement variance components	10
2.2 Measurement system capability metrics	11
2.2.1 Description of measurement system capability metrics	11
2.2.2 Research on measurement system capability metrics	12
2.3 Research gap	13
Chapter 3 Modeling	15
3.1 Capability criteria	16
3.1.1 Measurement system capability metrics	16
3.1.2 Process capability metrics	18
3.2 MLS confidence intervals	19

3.3	Balanced two-factor crossed random model with interaction	21
3.3.1	The model	21
3.3.2	MLS confidence intervals for variance components	23
3.3.3	MLS confidence intervals for measurement system capability metrics .	26
3.3.4	MLS confidence intervals for process capability metrics	29
3.4	Balanced two-factor crossed random model without interaction	29
3.4.1	The model	30
3.4.2	MLS confidence intervals for variance components	31
3.4.3	MLS confidence intervals for measurement system capability metrics .	34
3.4.4	MLS confidence intervals for process capability metrics	36
3.5	Algorithm	36
3.5.1	Assumptions and limitations	36
3.5.2	Procedure	37

Chapter 4 Sensitivity analysis of variability components and capability criteria **40**

4.1	Analysis of variance components	41
4.1.1	Analysis of repeatability σ_{Rep}^2	42
4.1.2	Analysis of part variability σ_p^2	45
4.1.3	Analysis of operator variability σ_o^2	48
4.1.4	Analysis of operator by part variability σ_{op}^2	52
4.1.5	Analysis of reproducibility $\sigma_{Reproducibility}^2$	56
4.1.6	Analysis of total measurement system variability σ_{GRR}^2	58
4.1.7	Analysis of total variability σ_{Total}^2	62

4.2	Analysis of measurement system capability metrics	65
4.2.1	Analysis of part variability to gauge variability ratio ρ_p	66
4.2.2	Analysis of proportion of measurement system variability due to re- peatability ρ_{Rep}	69
4.2.3	Analysis of proportion of measurement system variability due to repro- ducibility $\rho_{Reproducibility}$	73
4.2.4	Analysis of precision-to-tolerance ratio P/T	76
4.2.5	Analysis of signal-to-noise ratio SNR	79
4.3	Analysis of process capability metrics	82
4.3.1	Analysis of observed potential capability of the process C_p	83
4.3.2	Analysis of actual potential capability of the process C_p^*	86
4.4	Comparison of results	89
Chapter 5 Case study		92
5.1	General background	92
5.2	Gauge R&R study design	93
5.3	Data validation	94
5.4	Contrast of scenarios	97
Chapter 6 Conclusions		100
6.1	Summary	100
6.2	Contributions of the thesis	100
6.3	Future research	101
References		102

Appendices	109
A Numerical simulation data example	109
B Off-centering error data	123
C ANOVA tables	126

List of Tables

Table 3.1	P/T criteria	18
Table 3.2	SNR criteria	18
Table 3.3	C_p criteria	19
Table 3.4	ANOVA with interaction term	21
Table 3.5	ANOVA without interaction term	30
Table 4.1	Total measurement number for different parameter allocations	41
Table 4.2	Average CI width of σ_{Rep}^2 for different total measurement numbers	45
Table 4.3	Average CI width of σ_p^2 for different total measurement numbers	49
Table 4.4	Average CI width of σ_o^2 for different total measurement numbers	52
Table 4.5	Average CI width of σ_{op}^2 for different total measurement numbers	55
Table 4.6	Average CI width of $\sigma_{Reproducibility}^2$ for different total measurement numbers	59
Table 4.7	% reduction \overline{CI} width σ_{GRR}^2 for different designs with same measurement number	62
Table 4.8	% reduction \overline{CI} width σ_{Total}^2 for different designs with same measurement number	66
Table 4.9	Average CI width of ρ_p for different total measurement numbers	70
Table 4.10	Average CI width of ρ_{Rep} for different total measurement numbers	73
Table 4.11	Average CI width of $\rho_{Reproducibility}$ for different total measurement numbers	76
Table 4.12	Average CI width of P/T for different total measurement numbers	79
Table 4.13	Average CI width of SNR for different total measurement numbers	83
Table 4.14	Average CI width of C_p for different total measurement numbers	86

Table 4.15	Average CI width of C_p^* for different total measurement numbers . . .	89
Table 5.1	ρ_p results	98
Table 5.2	P/T results	98
Table 5.3	C_p^* results	99
Table A.1	Simulation data example for average confidence interval width of σ_{Rep}^2	109
Table A.2	Simulation data example for average confidence interval width of σ_p^2 .	110
Table A.3	Simulation data example for average confidence interval width of σ_o^2 .	111
Table A.4	Simulation data example for average confidence interval width of σ_{op}^2 .	112
Table A.5	Simulation data example for average confidence interval width of $\sigma_{Reproducibility}^2$	113
Table A.6	Simulation data example for average confidence interval width of σ_{GRR}^2	114
Table A.7	Simulation data example for average confidence interval width of σ_{Total}^2	115
Table A.8	Simulation data example for average confidence interval width of ρ_p . .	116
Table A.9	Simulation data example for average confidence interval width of ρ_{Rep}	117
Table A.10	Simulation data example for average confidence interval width of $\rho_{Reproducibility}$	118
Table A.11	Simulation data example for average confidence interval width of P/T	119
Table A.12	Simulation data example for average confidence interval width of SNR	120
Table A.13	Simulation data example for average confidence interval width of C_p .	121
Table A.14	Simulation data example for average confidence interval width of C_p^* .	122
Table B.15	Off-center error for experiment 1	123
Table B.16	Off-center error for experiment 2	124
Table B.17	Off-center error for experiment 3	125
Table C.18	ANOVA table for experiment 1	126
Table C.19	ANOVA table for experiment 2	126

Table C.20 ANOVA table for experiment 3 126

List of Figures

Figure 4.1	Average CI width of σ_{Rep}^2 for different sample size p	43
Figure 4.2	Average CI width of σ_{Rep}^2 for different sample size p	43
Figure 4.3	Average CI width of σ_{Rep}^2 for different replication n	44
Figure 4.4	Average CI width of σ_p^2 for different sample size p	46
Figure 4.5	Average CI width of σ_p^2 for different number of operator o	47
Figure 4.6	Average CI width of σ_p^2 for different replications n	48
Figure 4.7	Average CI width of σ_o^2 for different sample size p	49
Figure 4.8	Average CI width of σ_o^2 for different number of operator o	50
Figure 4.9	Average CI width of σ_o^2 for different replication n	51
Figure 4.10	Average CI width of σ_{op}^2 for different sample size p	53
Figure 4.11	Average CI width of σ_{op}^2 for different number of operator o	54
Figure 4.12	Average CI width of σ_{op}^2 for different replication n	54
Figure 4.13	Average CI width of $\sigma_{Reproducibility}^2$ for different sample size p	56
Figure 4.14	Average CI width of $\sigma_{Reproducibility}^2$ for different number of operator o	57
Figure 4.15	Average CI width of $\sigma_{Reproducibility}^2$ for different replication n	58
Figure 4.16	Average CI width of σ_{GRR}^2 for different sample size p	59
Figure 4.17	Average CI width of σ_{GRR}^2 for different number of operator o	60
Figure 4.18	Average CI width of σ_{GRR}^2 for different replication n	61
Figure 4.19	Average CI width of σ_{Total}^2 for different sample size p	63
Figure 4.20	Average CI width of σ_{Total}^2 for different number of operator o	64
Figure 4.21	Average CI width of σ_{Total}^2 for different replication n	65

Figure 4.22	Average CI width of ρ_p for different sample size p	67
Figure 4.23	Average CI width of ρ_p for different number of operator o	68
Figure 4.24	Average CI width of ρ_p for different replication n	69
Figure 4.25	Average CI width of ρ_{Rep} for different sample size p	71
Figure 4.26	Average CI width of ρ_{Rep} for different number of operator o	71
Figure 4.27	Average CI width of ρ_{Rep} for different replication n	72
Figure 4.28	Average CI width of $\rho_{Reproducibility}$ for different sample size p	74
Figure 4.29	Average CI width of $\rho_{Reproducibility}$ for different number of operator o	75
Figure 4.30	Average CI width of $\rho_{Reproducibility}$ for different replication n	75
Figure 4.31	Average CI width of P/T for different sample size p	77
Figure 4.32	Average CI width of P/T for different number of operator o	78
Figure 4.33	Average CI width of P/T for different replication n	78
Figure 4.34	Average CI width of SNR for different sample size p	80
Figure 4.35	Average CI width of SNR for different number of operator o	81
Figure 4.36	Average CI width of SNR for different replication n	82
Figure 4.37	Average CI width of C_p for different sample size p	84
Figure 4.38	Average CI width of C_p for different number of operator o	84
Figure 4.39	Average CI width of C_p for different replication n	85
Figure 4.40	Average CI width of C_p^* for different sample size p	87
Figure 4.41	Average CI width of C_p^* for different number of operator o	88
Figure 4.42	Average CI width of C_p^* for different replication n	88
Figure 4.43	Effect of sample size on average confidence interval width	90
Figure 4.44	Effect of number of operators on average confidence interval width	90

Figure 4.45	Effect of replications on average confidence interval width	91
Figure 5.1	Manufactured metal piece	93
Figure 5.2	Probability plot for normal distribution for experiment 1	94
Figure 5.3	Probability plot for normal distribution for experiment 2	95
Figure 5.4	Probability plot for normal distribution for experiment 3	95
Figure 5.5	$\bar{x} - R$ chart for experiment 1	96
Figure 5.6	$\bar{x} - R$ chart for experiment 2	96
Figure 5.7	$\bar{x} - R$ chart for experiment 3	97

Nomenclature

ANOVA	Analysis of variance
CI	Confidence interval
C_p	Observed potential capability of the process
C_p^*	Actual potential capability of the process
L	Confidence interval lower bound
LSL	Lower specification limit
MSA	Measurement system analysis
MLS	Modified-large-sample
n	Number of replications
o	Number of operators
p	Sample size
P/T	Precision-to-tolerance ratio
PCI	Process capability index
R&R	Repeatability and reproducibility
SNR	Signal-to-noise ratio
U	Confidence interval upper bound
USL	Upper specification limit
ρ_p	Part variability to gauge variability ratio
ρ_{Rep}	Proportion of measurement system variability due to repeatability
$\rho_{Reproducibility}$	Proportion of measurement system variability due to reproducibility
σ_{GRR}^2	Total measurement system variability

σ_o^2	Operator variability
σ_{op}^2	Operator by part variability
σ_p^2	Part variability
σ_{Rep}^2	Repeatability
$\sigma_{Reproducibility}^2$	Reproducibility
σ_{Total}^2	Total variability

Chapter 1

Introduction

1.1 Motivation

Manufacturing environment relies on fulfillment of product specifications and customer satisfaction. Quality characteristics of products are measured for assessing its accordance to requirements, but measurement error might mask true process performance (Harry and Lawson, 1992). Measurement system variability must be small with respect to product specifications as well to process variation. Hence, product evaluation and process improvement require reliable measurement techniques (Smith et al., 2007).

Total variation in a manufacturing process is a combination of part-to-part variation and measurement variation (Lin et al., 1997). A measurement system is useful only if it can detect changes in the process (Ingram and Taylor, 1998). Thus, prior to endeavor in process analysis and enhancement it is highly recommended to perform a measurement system capability study (Senvar and Firat, 2010) (Al-Refaie and Bata , 2010). It is carried out for assessing adequacy of the measurement instrument, isolate sources of variability and estimate total variability reliance on gauge variability (Burdick et al., 2003). The analysis of the measurement system has to take into account at least three basic components, appraiser, equipment and part (Ramesh and Sarma, 2013). Depending on the allocation of these parameters, conclusions with respect to the quality of the measurement system can highly vary. Therefore, the importance of a correct measurement system capability study design.

1.2 Objectives of the thesis

Main objective of this thesis is to analyze how sensitive are the confidence intervals of variability components and capability criteria to the design of a Gauge R&R study. Findings will be of help as guidance to practitioners when dealing with parameter allocation decisions in a Gauge R&R study. By following suggestions of the present research, it is expected to obtain more useful conclusions about the adequacy of a measurement system.

1.3 Research in the thesis

Measurement system analysis (MSA) relies on statistical tools, such as Gauge R&R study, to ensure that the measurement system is in acceptable conditions to monitor the manufacturing process (Sivaji, 2006). MSA considers metrics to estimate capability of the measurement system. Woodall and Borror (2008) classified them as those that compare measurement error to total or part error and those that compare measurement error to tolerance band.

In addition to point estimates, confidence intervals (CI) of measurement variance components can be also determined. In fact, these confidence intervals give more statistical information than point estimates (Montgomery and Runger, 1993a). Due to uncertainty of estimates, intervals enable to detect how small and how large variance components may really be (Wang and Li, 2003).

In this thesis, numerical simulation is considered for performing an sensitivity analysis of how the design of an Gauge R&R study affects the confidence interval width of variability components and capability criteria commonly used in industry to evaluate the measurement system.

1.4 Thesis structure

Literature review summarizing the basic principles of measurement system analysis is presented in Chapter 2. Based on previous research publications, point estimators and confidence interval calculations for variability components and capability criteria are introduced in Chapter 3. At the end of the chapter is presented the algorithm used for estimating the width of CIs. Chapter 4 includes the simulation results obtained from the algorithm and a discussion concerning the sensitivity analysis performed. An example problem is given in Chapter 5 to illustrate findings from previous chapters. Conclusions and future research are presented in Chapter 6.

Chapter 2

Literature review

2.1 Gauge repeatability and reproducibility study

2.1.1 Introduction

Raffaldi and Kappel (2004, pp. 1) stated that *"if measurement variation can be reduced and GR&R ratios improved, it is easier to differentiate between parts that are in or out of specifications, allowing parts to be accepted or rejected with greater confidence"*. Automotive Industry Action Group (AIAG, 2002) separated causes of measurement variation into location variation and width variation. Location variation includes accuracy or bias defined as the difference of the observed measurement with respect to a reference value, stability as a change in bias over time and linearity as a change in bias over the range of measurements. Width variation is divided into precision understood as proximity of observed measurements between each other, repeatability as the variation in consecutive measurements under same conditions and reproducibility as the variation in successive measurements under different conditions. Gauge repeatability and reproducibility (Gauge R&R) study is among the statistical approaches proposed to calculate measurement variability components (Al-Refaie and Bata , 2010). It is a powerful tool used to audit and improve measurement systems (Dasgupta and Murthy, 2001). Commonly Gauge R&R study considers repeatability as the variation observed when a part is measured several times using a measurement instrument under constant conditions, and reproducibility as the variation observed when a part is measured several times by different operators using the same gauge (Morris and Watson, 1998)(AIAG, 2002)(Burdick et al., 2003)(Smith et al., 2007)(Senvar and Oktay , 2010)(Senvar and Firat,

2010)(Antony et al., 1999)(Ramesh and Sarma, 2013)(Borror et al. , 1997). Repeatability is also known as equipment variation or within operator error, and reproducibility as appraiser variation or between operator error.

Depending on the results of gauge R&R study different authors have signaled corrective actions to improve the measurement system (Li and Al-Refaie, 2008). Senvar and Firat (2010) indicated that when repeatability is larger than reproducibility, it is a signal that the measurement instrument needs proper maintenance or to be redesigned for the intended use. Conversely, if reproducibility is larger, measurers need better training in how to use the measurement instrument. Morris and Watson (1998) demonstrated that re-training activities of personnel contributes to set in control variability in the R-chart and improves average results in the $\bar{x} - chart$. He et al. (2000) commented different strategies for improving measurement system capability depending on the source of measurement system variation affecting the process.

2.1.2 Gauge R&R study and its applications

Gauge R&R study has been conducted to assess measurement system adequacy in industry sectors. In automobile industry, Tsai (1988) presented an example of how to evaluate a measurement tool utilized to inspect an inject molding part during production. Measurers and parts were randomly selected from the process to perform a gauge R&R study and results showed that the instrument had the necessary capability and no corrective actions were required. Osma (2011) compared two common methodologies for performing a gauge R&R study. It was considered real cases from 4 vehicle components: a stabilizer clamp bracket, a coated steel sheet, a steel sheet body panel, and a brake disc. The author set out

a procedure on how to select the proper methodology for different conditions of the study.

In pharmaceutical industry, Dejaegher et al. (2006) used Gauge R&R study to compare gauge capability of two laboratory tests employed to confirm presence of active pharmaceutical ingredients in drug substances. Authors established the factors that had the most significant effects on the capability of the high performance liquid chromatographic (HPLC) measurement method and suggested corrective actions. With the improved system, they achieved a performance similar to that by using the volumetric titration measurement method. Gao et al. (2007) set up a gauge R&R study for analyzing variability of a dissolution testing procedure. Authors examined contribution of instrument, operators and tablets variabilities to total measurement error. Based on the results a change in the gauge was made to enhance its performance.

For wood industry, Li and Al-Refaie (2008) performed an assay of the measurement system utilized for testing quality of wooden parts. By way of a Gauge R&R study it was established that the gauge system was unreliable for the intended application and corrective actions like re-training of workers, proper selection of tools and changes in the measurement procedure were proposed. A second Gauge R&R study was conducted after improving the process and results showed an increment on measurement system capability. Ile and Lazea (2014) presented a Gauge R&R study for establishing adequacy of a gauge. It was conducted in two phases. In the first phase 3 testers measured 10 wood items 3 times in a random order. Results indicated that the gauge was not appropriate. In a second phase the experiment was done using a more precise measurement tool and it was found a considerable amelioration of the measurement system.

In electronic industry, Houf and Berman (1988) investigated the adequacy of a power mod-

ule thermal test equipment. From Gauge R&R study it was inferred that the measurement system variability was unacceptable for the application and three alternatives were suggested to reduce its error. Wang (2008)) investigated the adequacy of a measurement system employed to test a semiconductor manufacturing process. Gauge R&R study results were not satisfactory and affected process yield estimation. The author pointed out that it was needed to reduce gauge variability in order to overcome this issue. Kaija et al. (2010) conducted a Gauge R&R study prior to process analysis. Conformance assessment of two measurement systems used for testing an inkjetting dielectric layer was carried into effect. Both studies established that the total gauge system variance was low in comparison to part-to-part error and no changes were required in the measurement systems.

For food industry, Srikaeo et al. (2005) conducted a research in a commercial wheat-based biscuit cooking process. Measurement system was evaluated through a Gauge R&R study. The study revealed a high contribution of gauge system variation to total error. Corrective actions such as re-training of appraisers and equipment recalibration were proposed to achieve adequacy of the measurement system.

2.1.3 Gauge R&R study design

With respect to planning a Gauge R&R study, several authors have suggested different arrangements for obtaining reliable conclusions. Having more personnel, parts and replications in the study improves confidence in the results but increases the experimental cost and time required. Whereby, it is needed to carefully design the Gauge R&R study (Tsai, 1988)(Pan, 2004). Antony et al. (1999) recommended 2 to 3 appraisers, 5 to 20 samples and at least 2 replications. He et al. (2000) advised the need of at least 2 measurers, 10 to 15 parts and

no less than 2 trials. (Pan, 2004) proposed a method using the shortest CI to decide the optimal allocation of number of personnel, parts and trials for a Gauge R&R study. Through simulation it was observed the impact of the different combinations of these parameters and distinct values of sources of measurement system variation on the width of σ_M^2 confidence interval. Zappa and Deldossi (2009) presented a procedure to select the combination of number of workers, parts and samples using misclassification rates.

Depending on measurement system conditions, different Gauge R&R study approaches have been proposed, like crossed study in which all parts are measured by all operators, nested study where each part is measured only by one appraiser without operator study (He et al., 2000).

2.1.4 Gauge R&R study methods

Average & range ($\bar{X} - R$) and analysis of variance (ANOVA) are the common methods used to carry out a Gauge R&R study (He et al., 2000)(Peruchi et al., 2013)(Osma, 2011). Wheeler and Lyday (1989) proposed techniques to evaluate the measurement process through $\bar{X} - R$ control charts. Montgomery and Runger (1993b) presented a methodology to estimate measurement variance components through ANOVA analysis.

$\bar{X} - R$ method utilizes range and bias correction factor d_2 for estimating variance components. Interpretation of these charts used in a Gauge R&R study differs from charts employed in monitoring the behavior of a manufacturing process. Under this context the average chart indicates the ability of the measurement system to discern between parts while the range chart shows the magnitude of the measurement error (Montgomery, 2012).

Morris and Watson (1998) compared results of $\bar{X} - R$ and individual & moving range (X-MR)

control charts and demonstrated that latter might yield to an inflated value of measurement variability. Despite its relative easiness to estimate measurement error, $\bar{X} - R$ method has some drawbacks. It can be applied only to two-factor studies and does not allow confidence intervals estimation (Senvar and Oktay , 2010). This approach only considers gauge and appraiser contribution to variability but omits operator-by-part interaction effect (Antony et al., 1999)(He et al., 2000). Montgomery and Runger (1993b) showed that $\bar{X} - R$ method can substantially underestimate measurement error specially when operator-by-part interaction effect is large. Modifications on formulas for calculating variance components have arisen over the years, especially on reproducibility component (AIAG, 2002)(Wheeler, 2009)(Ramesh and Sarma, 2013).

Using ANOVA approach, Gauge R&R study can be performed resembling a two-factor design of experiments (DOE) model that includes factors, levels and replicates (Antony et al., 1999). The effect of appraiser, sample, appraiser-sample interaction and replication over the observed measurement response is analyzed in this model (Montgomery and Runger, 1993b).

ANOVA method allows point and interval estimation of variance components, enables calculation of operator-by-part interaction effect and can be applied to complex measurement system analysis (Burdick et al., 2003)(Pan, 2006). Although ANOVA method offers significant advantages, it might produce negative estimation of variance components that would seriously underestimate total measurement variability. Montgomery and Runger (1993b) proposed a modified analysis of variance method to avoid calculation of negative components. In addition to ANOVA method, there are other useful approaches. Maximum likelihood procedure and MINQUE procedure are among them. Montgomery and Runger (1993a) signaled that latter two approaches guarantee calculation of non-negative variance components and

produce shorter confidence intervals than ANOVA method.

2.1.5 Confidence intervals for measurement variance components

Borror et al. (1997) proposed calculations of CIs for repeatability, reproducibility and total gauge variation components by using restricted maximum likelihood (RELM) and modified large-sample (MLS) methods. It showed that RELM method estimates shorter intervals than MLS method when few operators are considered in the study. By increasing the number of appraisers it obtained similar interval widths from both methods. Burdick and Larsen (1997) conducted a simulation study for comparing the confidence coefficient and estimated average CI widths based on Satterthwaite (SATT), Automotive Industry Action Group (AIAG), REML and MLS methods. Among them, MLS method was recommended since it is more reliable under different experimental conditions. Wang and Li (2003) estimated CIs using Bootstrap method and variance components calculated with $\bar{X} - R$ approach. It was found similar results when comparing CIs obtained from maximum likelihood (MLE), restricted maximum likelihood (RMLE) and MLS methods determined using ANOVA approach. Park and Yoon (2009) evaluated through simulation suitability of confidence intervals for repeatability, reproducibility and total measurement error based on MLS method and generalized p-value (GEN) method. Results suggested that both methods maintain the expected confidence coefficient and similar interval widths.

Montgomery and Runger (1993b) noticed that increasing the number of replications has a higher impact on reducing the width of repeatability CI than on part or operator CI. They also pointed out that using many parts in an experiment increases the likelihood for variation's sources to occur when compared to repeated measurements on the same sample. Burdick

and Larsen (1997) demonstrated that when performing a Gauge R&R capability study, the configuration of the experiment affects the width of the CI. It was established that whether the appraiser variance component is of interest, it is necessary to include sufficient number of operators in the study to calculate a shorter CI. On the other hand, if part variance component is the main purpose for the analysis, a design of experiment that maximizes the number of samples provides CI that is short enough to be of use. Increasing number of replicates showed a slightly reduction of CI width, thus it was recommended to reduce the number of replications and to increase either the number of samples or the number of operators.

2.2 Measurement system capability metrics

2.2.1 Description of measurement system capability metrics

Total gauge R&R %Study Variance (%R&R) evaluates the ratio $\hat{\sigma}_M/\hat{\sigma}_T$. AIAG (2002) considered that a value of %R&R less than 0.1 is acceptable, between 0.1 and 0.3 might be acceptable depending on measurement system conditions, and greater than 0.3 is deemed to be unacceptable. Precision-to-tolerance ratio (P/T) estimates gauge capability with respect to specification limits. Values less than or equal to 0.1 imply that the measurement system is adequate. Mader et al. (1999) noticed that P/T can be a misleading criterion to assess gauge capability because different levels of measurement system P/T can be tolerated depending on process capability. Signal-to-noise ratio (SNR) is a function of σ_P^2/σ_M^2 . AIAG (2002) established that a value of SNR equal to or greater than 4.0 denotes that the measurement system is appropriate for the corresponding application, and less than 2.0 means that the measurement system is of no value for the intended use. Discrimination ratio (DR) is another

function of σ_P^2/σ_M^2 that estimates how many categories are distinguished by the measurement system. AIAG (2002) stated that a measurement system is capable if DR exceeds 4.0. In addition to these metrics, process capability indices (PCIs), such as potential process capability (C_p) and actual process capability (C_{pk}) have been employed to validate how measurement system error affects estimation of process performance (Lin et al., 1997).

2.2.2 Research on measurement system capability metrics

Several authors have discussed how these metrics have improved after corrective actions were implemented over the measurement system. Lin et al. (1997) used an algorithm based on Taguchi robust design and Gauge R&R study to evaluate and reduce error of measurement system. It was shown how these initiatives had a positive effect on measurement system by reducing %R&R, likewise on process performance by increasing C_p and C_{pk} . Bordignon and Scagliarini (2002) analyzed the effect of measurement error on process capability indices C_p and C_{pk} of a sample. They noticed that the sampling error has an overestimating effect on PCIs, conversely the measurement error has an underestimating effect on PCIs. Majeske and Andrews (2002) utilized P/T, C_p and correlation on repeated measurements to propose a methodology for evaluating measurement and manufacturing processes between supplier and customer. Larsen (2003) conducted a Gauge R&R study using ANOVA method. P/T, SNR and C_{pk} were calculated for assessing adequacy of the measurement system and manufacturing process. Confidence intervals of variance components and metrics were also calculated for verifying the precision of these estimates. Pearn and Liao (2005) conducted an analysis to evaluate impact of measurement error on C_{pk} index. The analysis revealed that this index severely underestimate true process capability when measurement error is present.

Authors recommended the use of modified confidence bounds and critical values to deal with data contaminated with measurement error. Hsu et al. (2007) performed an investigation to analyze the performance of C_{pmk} and its lower confidence bound when the sample data is affected by measurement error. Results demonstrated that gauge variability reduces estimation of true process capability. The authors proposed 3 methods for adjusting the lower confidence bound of this PCI. Pearn and Liao (2007) studied the effect of measurement error over process capability index C_p and its CI. It was established that this error originates a decrement on α -risk and power of test. Authors proposed adjusted CI bounds and critical values of C_p to account for measurement variability. Li and Al-Refaie (2008) used Six Sigma methodology and Gauge R&R study to investigate and reach suitability of the measurement system. They reported improvement of %R&R, P/T and SNR after corrective actions were implemented. Al-Refaie and Bata (2010) presented a methodology for assessing capability of measurement system and manufacturing process through the analysis of P/T, SNR, DR, C_p and C_{pk} . Al-Refaie (2011) considered the use of metrics calculated through $\bar{X} - R$ method for establishing two procedures for manufacturing and measurement capability assessment. The first method considered the use of P/T, %R&R and C_p for evaluating capability and the second P/T, number of distinct categories (NDC) and C_p . Effectiveness of both procedures was demonstrated through case studies.

2.3 Research gap

Previous research focused mainly on analysing how Gauge R&R study design affects the confidence interval width of the principal variability components of measurement error. However, in a real manufacturing environment, decisions about the quality of the measurement system

are not taken with respect to the measurement system variability components, but with respect to the measurement system capability metrics. Therefore the importance of analysing parameter allocation impact on the reduction of the CI width of these metrics.

Chapter 3

Modeling

Regardless of process characteristics, any measurement activity is affected by certain level of variability. This phenomenon can be expressed as:

$$Y = X + \varepsilon. \quad (3.1)$$

where Y represents the observed measurement value by the appraiser on a randomly selected sample, X is the real measurement value of the sample and ε the error due to the measurement activity (Burdick et al., 2003). Y , X and ε are, in many manufacturing processes, independent normal random variables. We denote the mean and variance of these 3 random variables by μ_{Total} , μ_{Part} , μ_{GRR} and σ_{Total}^2 , σ_{Part}^2 , σ_{GRR}^2 , respectively. It is commonly assumed that μ_{GRR} can be eliminated through a correct calibration of the measurement system. Measurement system variability σ_{GRR}^2 can be decomposed into two sources of variation, Repeatability and reproducibility (AIAG, 2002), denoted as σ_{Rep}^2 and $\sigma_{Reproducibility}^2$ respectively. Repeatability represents the variability observed when an operator measures a sample several times using the same gauge, and reproducibility the variability originated by different operators when measuring several times a part with the same gauge.

Conformity of a part to process requirements can be wrongly judged by cause of an error in the measurement activity. Capability assessment of a measurement system requires a way to interpret and minimize the sources of measurement error, due to operator or gauge, or both. This can be achieved by analyzing different capability parameters. In industry, population parameters for these criteria are not often at hand and it is common to approximate them

through point estimations. Due to uncertainty, it may be necessary to complement these sample statistics with their respective confidence intervals.

Width of the confidence interval is of interest when performing a measurement system analysis. Narrower CIs are desirable when concluding about the quality of the measurement system. If the CI is too wide, the estimation might not be of practical use for assessing capability of the measurement system. Depending on the composition of total variability and the allocation of resources in the Gauge R&R study, i.e. number of operators, samples, replications, etc., width of CIs can highly vary.

Along this chapter it is presented some of the capability criteria mentioned in literature and their respective Modified-large-sample (MLS) confidence intervals. Final section introduces the algorithm employed to calculate CI width of variability components and capability criterion point estimations for different levels of variability and resources used in a typical Gauge R&R study.

3.1 Capability criteria

In a Gauge R&R study several quality measures may be considered as capability of the measurement system and the monitored process. These parameters are functions of the variability components. This section includes a description of point estimations of measurement system capability metrics and process capability metrics.

3.1.1 Measurement system capability metrics

A Gauge R&R study aims to determine if measurement system variability is small compared to manufacturing process variability. Hence, it is of use to analyze the ratio of part variability

to gauge variability $\hat{\rho}_P$.

$$\hat{\rho}_P = \frac{\hat{\sigma}_P^2}{\hat{\sigma}_{GRR}^2} \quad (3.2)$$

It is also of interest to compare repeatability and reproducibility contribution to total measurement system variation. As a result of this analysis it is possible to decide the best approach for improving the measurement system. Therefore, it is useful to calculate the proportion of measurement system variability due to repeatability and reproducibility, denoted as $\hat{\rho}_{Repeatability}$ and $\hat{\rho}_{Reproducibility}$ respectively.

$$\hat{\rho}_{Repeatability} = \frac{\hat{\sigma}_{Rep}^2}{\hat{\sigma}_{GRR}^2} \quad (3.3)$$

$$\hat{\rho}_{Reproducibility} = \frac{\hat{\sigma}_{Reproducibility}^2}{\hat{\sigma}_{GRR}^2} \quad (3.4)$$

Precision-to-tolerance ratio ($\widehat{P/T}$) is a quality measure often employed for evaluating the measurement system. It determines the percentage of the tolerance band used by the measurement system variation. It is calculated as:

$$\widehat{P/T} = \frac{k\hat{\sigma}_{GRR}}{USL - LSL} \quad (3.5)$$

where USL and LSL are the upper and lower specification limits respectively, and k represents the number of standard deviations between the natural tolerance limits of a normal distributed process. Common values of k are 5.15 and 6 that accounts for 99% and 99.73% of measurement system variation respectively. AIAG (2002) evaluates the quality of the

measurement system according to the criteria found on Table 3.1

Table 3.1: P/T criteria

P/T	Measurement system
< 10%	Adequate
Between 10% and 30%	Moderate
> 30%	Inadequate

Signal-to-noise ratio (\widehat{SNR}) estimates the number of distinct categories of parts that a measurement system can distinguish. It can be expressed as:

$$\widehat{SNR} = \frac{2\sigma_P^2}{\sigma_{GRR}^2} \quad (3.6)$$

AIAG (2002) establishes the following criteria detailed on Table 3.2 for judging capability of the measurement system according to this metric

Table 3.2: SNR criteria

SNR	Measurement system
< 2%	Inadequate
Between 2% and 4%	Moderate
> 4%	Adequate

3.1.2 Process capability metrics

Metrics discussed in the previous section are intended to study quality of the measurement system, but it might also be convenient to estimate performance of the monitored process.

One common metric is the ratio \hat{C}_p that determines the observed potential capability of the process when its mean is centered in the tolerance band.

$$\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}_{Total}} \quad (3.7)$$

In general, criteria from Table 3.3 are used to evaluate quality of the monitored process.

Table 3.3: C_p criteria

C_p	Process quality
< 1	Not acceptable
Between 1 and 1.33	Moderate
> 1.33	Acceptable

Previous metric considers measurement error when it is calculated. Another useful capability metric that considers only part variability is the actual potential capability of the process \hat{C}_p^* .

$$\hat{C}_p^* = \frac{USL - LSL}{6\hat{\sigma}_p} \quad (3.8)$$

3.2 MLS confidence intervals

Data collected in a Gauge R&R study can be used to estimate population parameters. Confidence intervals help to deal with the sample uncertainty. In general, the $100(1 - \alpha)\%$ confidence interval of a parameter ω consists in a Lower and an Upper bound, denoted by L and U respectively, such that

$$Pr[L \leq \omega \leq U] = 1 - \alpha \quad (3.9)$$

Different methods have been proposed for constructing CIs for the variability components and capability criteria for point estimations discussed above. Among them, modified large-sample (MLS) method introduced by Graybill and Wang (1980) has been highly recommended in literature. MLS method first approximates a large-sample confidence interval for the parameter in study, and depending on its particular conditions, the CI is recalculated to make it

exact (Burdick et al., 2005a). Modifications of MLS confidence intervals have been proposed to analyze parameters under different conditions. Notations above are used to calculate MLS confidence intervals throughout following sections.

$$G_q = 1 - \frac{df_q}{\chi_{\frac{\alpha}{2}, df_q}^2} \quad (3.10)$$

$$H_q = \frac{df_q}{\chi_{1-\frac{\alpha}{2}, df_q}^2} - 1 \quad (3.11)$$

$$I_{q,r} = F_{\frac{\alpha}{2}, df_q, df_r} \quad (3.12)$$

$$J_{q,r} = F_{1-\frac{\alpha}{2}, df_q, df_r} \quad (3.13)$$

$$G_{q,r} = \frac{(F_{\frac{\alpha}{2}, df_q, df_r} - 1)^2 - G_q^2 F_{\frac{\alpha}{2}, df_q, df_r}^2 - H_r^2}{F_{\frac{\alpha}{2}, df_q, df_r}} \quad (3.14)$$

$$H_{q,r} = \frac{(1 - F_{1-\frac{\alpha}{2}, df_q, df_r})^2 - H_q^2 F_{1-\frac{\alpha}{2}, df_q, df_r}^2 - G_r^2}{F_{1-\frac{\alpha}{2}, df_q, df_r}} \quad (3.15)$$

$$G_{q,r}^* = \left(1 - \frac{df_q + df_r}{\chi_{\frac{\alpha}{2}, df_q + df_r}^2}\right)^2 \frac{(df_q + df_r)^2}{df_q df_r} - \frac{G_q^2 df_q}{df_r} - \frac{G_r^2 df_r}{df_q} \quad (3.16)$$

where $\chi_{\frac{\alpha}{2}, df_q}^2$ represents the chi-square value with df_q degrees of freedom and $\alpha/2$ area to the right, and $F_{\frac{\alpha}{2}, df_q, df_r}$ is the F -value with df_q and df_r degrees of freedom and $\alpha/2$ area to the right.

3.3 Balanced two-factor crossed random model with interaction

A classical Gauge R&R study considers the effects of two main factors, operator and measurement equipment, on total system variation. An experimental design approach can be of use to describe the gauge capability study in which variance components are estimated.

3.3.1 The model

In the balanced two-factor crossed random model with interaction above, X_{ijk} represents the measurement value obtained by appraiser i on sample j at replication k . Montgomery and Runger (1993b) presented the model

$$X_{ijk} = \mu + O_i + P_j + (OP)_{ij} + R_{k(ij)} \quad (3.17)$$

where $i = 1, \dots, o$; $j = 1, \dots, p$; $k = 1, \dots, n$

For this model, μ is the overall mean of all observations, terms O_i and P_j represent the main effects of operator and part factors on the measurement respectively, $(OP)_{ij}$ is the operator by part interaction effect on the measurement and $R_{k(ij)}$ denotes the replication effect on X_{ijk} . Independency, randomness and normality behavior are assumed for these effects, where $O_i \sim N(0, \sigma_o^2)$, $P_j \sim N(0, \sigma_p^2)$, $(OP)_{ij} \sim N(0, \sigma_{op}^2)$ and $R_{k(ij)} \sim N(0, \sigma_{Rep}^2)$. ANOVA table for this model is shown on Table 3.4.

Table 3.4: ANOVA with interaction term

Source of variation	Sum of Squares	Degrees of Freedom	Mean Square	Expected mean square	F_o	F-statistic
P - Part	SS_p	$df_p = p - 1$	$MS_p = SS_p/df_p$	$\theta_p = \sigma_{Rep}^2 + n\sigma_{op}^2 + on\sigma_p^2$	MS_p/MS_{op}	$F_{\alpha, df_p, df_{op}}$
O - Operator	SS_o	$df_o = o - 1$	$MS_o = SS_o/df_o$	$\theta_o = \sigma_{Rep}^2 + n\sigma_{op}^2 + pn\sigma_o^2$	MS_o/MS_{op}	$F_{\alpha, df_o, df_{op}}$
OP - Operator x Part	SS_{op}	$df_{op} = (p - 1)(o - 1)$	$MS_{op} = SS_{op}/df_{op}$	$\theta_{op} = \sigma_{Rep}^2 + n\sigma_{op}^2$	MS_{op}/MS_{Rep}	$F_{\alpha, df_{op}, df_{Rep}}$
R - Repeatability	SS_{Rep}	$df_{Rep} = po(n - 1)$	$MS_{Rep} = SS_{Rep}/df_{Rep}$	$\theta_{Rep} = \sigma_{Rep}^2$		
T - Total	SS_{Total}	$df_{Total} = pon - 1$				

Montgomery and Runger (1993a) defined the point estimations of the variance components

as:

$$\hat{\sigma}_{Rep}^2 = MS_{Rep} \quad (3.18)$$

$$\hat{\sigma}_o^2 = \frac{MS_o - MS_{op}}{pn} \quad (3.19)$$

$$\hat{\sigma}_{op}^2 = \frac{MS_{op} - MS_{Rep}}{n} \quad (3.20)$$

$$\hat{\sigma}_p^2 = \frac{MS_p - MS_{op}}{on} \quad (3.21)$$

Reproducibility is calculated as the sum of the variation due to operator effect and the interaction between operator and part effect, denoted as:

$$\hat{\sigma}_{Reproducibility}^2 = \hat{\sigma}_o^2 + \hat{\sigma}_{op}^2 = \frac{MS_o + (p-1)MS_{op} - pMS_{Rep}}{pn} \quad (3.22)$$

Measurement system variation is attributed to all variability components except part variation

$$\hat{\sigma}_{GRR}^2 = \hat{\sigma}_{Rep}^2 + \hat{\sigma}_{Reproducibility}^2 = \frac{MS_o + (p-1)MS_{op} - p(n-1)MS_{Rep}}{pn} \quad (3.23)$$

and total variation is the sum of all sources of variation in the system:

$$\hat{\sigma}_{Total}^2 = \hat{\sigma}_p^2 + \hat{\sigma}_{GRR}^2 = \frac{pMS_p + oMS_o + [o(p-1) - p]MS_{op} + op(n-1)MS_{Rep}}{pon} \quad (3.24)$$

3.3.2 MLS confidence intervals for variance components

Repeatability can be expressed as $\sigma_{Rep}^2 = \theta_{Rep}$. Burdick and Larsen (1997) presented an exact MLS confidence interval for σ_{Rep}^2 .

$$(1 - G_{Rep})MS_{Rep} \leq \sigma_{Rep}^2 \leq (1 + H_{Rep})MS_{Rep} \quad (3.25)$$

Variability of operator σ_o^2 , interaction between operator and part variation σ_{op}^2 , manufacturing process variability σ_p^2 and reproducibility can be expressed as $\sigma_o^2 = (\theta_o - \theta_{op})/(pn)$, $\sigma_{op}^2 = (\theta_{op} - \theta_{Rep})/n$, $\sigma_p^2 = (\theta_p - \theta_{op})/(on)$ and $\sigma_{Reproducibility}^2 = [\theta_o + (p - 1)\theta_{op} - p\theta_{Rep}]/(pn)$ respectively. Ting et al. (1990) proposed a MLS method for calculating the CI of linear combinations of expected mean squares that contains difference. Applying this method, CI of σ_o^2 is defined as

$$\hat{\sigma}_o^2 - \sqrt{VL_o} \leq \sigma_o^2 \leq \hat{\sigma}_o^2 + \sqrt{VU_o} \quad (3.26)$$

where

$$VL_o = \frac{G_o^2 MS_o^2 + H_{op}^2 MS_{op}^2 + G_{o,op} MS_o MS_{op}}{(pn)^2} \quad (3.27)$$

$$VU_o = \frac{H_o^2 MS_o^2 + G_{op}^2 MS_{op}^2 + H_{o,op} MS_o MS_{op}}{(pn)^2} \quad (3.28)$$

CI of σ_{op}^2 is represented as

$$\hat{\sigma}_{op}^2 - \sqrt{VL_{op}} \leq \sigma_{op}^2 \leq \hat{\sigma}_{op}^2 + \sqrt{VU_{op}} \quad (3.29)$$

where

$$VL_{op} = \frac{G_{op}^2 MS_{op}^2 + H_{Rep}^2 MS_{Rep}^2 + G_{op,Rep} MS_{op} MS_{Rep}}{n^2} \quad (3.30)$$

$$VU_{op} = \frac{H_{op}^2 MS_{op}^2 + G_{Rep}^2 MS_{Rep}^2 + H_{op,Rep} MS_{op} MS_{Rep}}{n^2} \quad (3.31)$$

CI of σ_p^2 is defined as

$$\hat{\sigma}_p^2 - \sqrt{VL_p} \leq \sigma_p^2 \leq \hat{\sigma}_p^2 + \sqrt{VU_p} \quad (3.32)$$

where

$$VL_p = \frac{G_p^2 MS_p^2 + H_{op}^2 MS_{op}^2 + G_{p,op} MS_p MS_{op}}{(on)^2} \quad (3.33)$$

$$VU_p = \frac{H_p^2 MS_p^2 + G_{op}^2 MS_{op}^2 + H_{p,op} MS_p MS_{op}}{(on)^2} \quad (3.34)$$

and CI of $\sigma_{Reproducibility}^2$ is expressed as

$$\hat{\sigma}_{Reproducibility}^2 - \sqrt{VL_{Reproducibility}} \leq \sigma_{Reproducibility}^2 \leq \hat{\sigma}_{Reproducibility}^2 + \sqrt{VU_{Reproducibility}} \quad (3.35)$$

where

$$\begin{aligned}
VL_{Reproducibility} = & [G_o^2 MS_o^2 + G_{op}^2 (p-1)^2 MS_{op}^2 + H_{Rep}^2 p^2 MS_{Rep}^2 \\
& + G_{o,Rep} p MS_o MS_{Rep} + G_{op,Rep} (p-1) p MS_{op} MS_{Rep} \\
& + G_{o,op}^* (p-1) MS_o MS_{op}] / (pn)^2
\end{aligned} \tag{3.36}$$

$$\begin{aligned}
VU_{Reproducibility} = & H_o^2 MS_o^2 + H_{op}^2 (p-1)^2 MS_{op}^2 + G_{Rep}^2 p^2 MS_{Rep}^2 \\
& + H_{o,Rep} p MS_o MS_{Rep} \\
& + H_{op,Rep} (p-1) p MS_{op} MS_{Repeatability}] / (pn)^2
\end{aligned} \tag{3.37}$$

Total measurement system variability σ_{GRR}^2 and total system variability σ_{Total}^2 can be expressed as $\sigma_{GRR}^2 = [\theta_o + (p-1)\theta_{op} + p(n-1)\theta_{Rep}] / (pn)$ and $\sigma_{Total}^2 = [p\theta_p + o\theta_o + (po - p - o)\theta_{op} + p(on - o)\theta_{Rep}] / (pon)$ respectively. Graybill and Wang (1980) discussed a variation to the MLS method for calculating the addition of expected mean square values. Using this method CI of σ_{GRR}^2 is defined as

$$\hat{\sigma}_{GRR}^2 - \sqrt{VL_{GRR}} \leq \sigma_{GRR}^2 \leq \hat{\sigma}_{GRR}^2 + \sqrt{VU_{GRR}} \tag{3.38}$$

where

$$VL_{GRR} = \frac{G_o^2 MS_o^2 + G_{op}^2 (p-1)^2 MS_{op}^2 + G_{Rep}^2 p^2 (n-1)^2 MS_{Rep}^2}{(pn)^2} \tag{3.39}$$

$$VU_{GRR} = \frac{H_o^2 MS_o^2 + H_{op}^2 (p-1)^2 MS_{op}^2 + H_{Rep}^2 p^2 (n-1)^2 MS_{Rep}^2}{(pn)^2} \quad (3.40)$$

and CI of σ_{Total}^2 is represented as

$$\hat{\sigma}_{Total}^2 - \sqrt{VL_{Total}} \leq \sigma_{Total}^2 \leq \hat{\sigma}_{Total}^2 + \sqrt{VU_{Total}} \quad (3.41)$$

where

$$\begin{aligned} VL_{Total} = & [G_p^2 p^2 MS_p^2 + G_o^2 o^2 MS_o^2 + G_{op}^2 (po - p - o)^2 MS_{op}^2 \\ & + G_{Rep}^2 p^2 o^2 (n-1)^2 MS_{Rep}^2] / (pon)^2 \end{aligned} \quad (3.42)$$

$$\begin{aligned} VU_{Total} = & [H_p^2 p^2 MS_p^2 + H_o^2 o^2 MS_o^2 + H_{op}^2 (po - p - o)^2 MS_{op}^2 \\ & + H_{Rep}^2 p^2 o^2 (n-1)^2 MS_{Rep}^2] / (pon)^2 \end{aligned} \quad (3.43)$$

3.3.3 MLS confidence intervals for measurement system capability metrics

Leiva and Graybill (1986) proposed a modification of MLS method to calculate exact confidence intervals of the ratio of part variability to total measurement system variability ρ_p .

$$L_{\rho_p} \leq \rho_p \leq U_{\rho_p} \quad (3.44)$$

where

$$L_{\rho_p} = \frac{p(1 - G_p)(MS_p - I_{p,op}MS_{op})}{po(n - 1)MS_{Rep} + o(1 - G_p)I_{p,o}MS_o + o(p - 1)MS_{po}} \quad (3.45)$$

$$U_{\rho_p} = \frac{p(1 + H_p)(MS_p - J_{p,op}MS_{op})}{po(n - 1)MS_{Rep} + o(1 + H_p)J_{p,o}MS_o + o(p - 1)MS_{po}} \quad (3.46)$$

Confidence interval for $\rho_{Repeatability}$ is presented by Burdick et al. (2005a). Authors expressed $\rho_{Repeatability} = 1/(\eta + 1)$, where $\eta = \sigma_{Reproducibility}^2/\sigma_{Rep}^2$. The value of η can be presented as

$$\eta = \frac{\theta_o + (p - 1)\theta_{po}}{pn\theta_{Rep}} + \frac{1}{n} \quad (3.47)$$

Considering a previous research done by Lu et al. (1987), CI for η can be obtained and derived for calculating the CI of ρ_{Rep} .

$$\frac{1}{U_{\rho_{Repeatability}}^* + 1} \leq \rho_{Repeatability} \leq \frac{1}{L_{\rho_{Repeatability}}^* + 1} \quad (3.48)$$

where

$$L_{\rho_{Repeatability}}^* = \frac{1}{pnMS_{Rep}} \left[\left(1 - \frac{2}{po(n - 1)} \right) (MS_o + (p - 1)MS_{po}) - \sqrt{V_{L_{Rep}}} \right] - \frac{1}{n} \quad (3.49)$$

$$U_{\rho_{Repeatability}}^* = \frac{1}{pnMS_{Rep}} \left[\left(1 - \frac{2}{po(n - 1)} \right) (MS_o + (p - 1)MS_{po}) + \sqrt{V_{U_{Rep}}} \right] - \frac{1}{n} \quad (3.50)$$

$$V_{LRep} = aMS_o^2 + b(p-1)^2MS_{op}^2 + c(p-1)MS_oMS_{op} \quad (3.51)$$

$$V_{URep} = dMS_o^2 + e(p-1)^2MS_{op}^2 + f(p-1)MS_oMS_{op} \quad (3.52)$$

$$a = \left[1 - \frac{2}{po(n-1)} - \frac{1}{I_{o,Rep}} \right]^2 \quad (3.53)$$

$$b = \left[1 - \frac{2}{po(n-1)} - \frac{1}{I_{po,Rep}} \right]^2 \quad (3.54)$$

$$c = \left[1 - \frac{2}{po(n-1)} - \frac{1}{F_{\frac{\alpha}{2}, pdf_o, df_{Rep}}} \right]^2 \frac{p^2}{p-1} - \frac{a}{p-1} - (p-1)b \quad (3.55)$$

$$d = \left[\frac{1}{J_{o,Rep}} - 1 + \frac{2}{po(n-1)} \right]^2 \quad (3.56)$$

$$e = \left[\frac{1}{J_{po,Rep}} - 1 + \frac{2}{po(n-1)} \right]^2 \quad (3.57)$$

$$f = \left[\frac{1}{F_{1-\frac{\alpha}{2}, pdf_o, df_{Rep}}} - 1 + \frac{2}{po(n-1)} \right]^2 \frac{p^2}{p-1} - \frac{d}{p-1} - (p-1)e \quad (3.58)$$

Ratio $\rho_{Reproducibility} = 1 - \rho_{Repeatability}$. From previous result, the confidence interval for

$\rho_{Reproducibility}$ can be defined as

$$1 - \frac{1}{L_{\rho_{Repeatability}}^* + 1} \leq \rho_{Reproducibility} \leq 1 - \frac{1}{U_{\rho_{Repeatability}}^* + 1} \quad (3.59)$$

Precision-to-tolerance P/T ratio is a function of total measurement system variability σ_{GRR}^2 .

The MLS confidence interval for P/T is expressed as

$$\frac{6\sqrt{\hat{\sigma}_{GRR}^2 - \sqrt{VL_{GRR}}}}{USL - LSL} \leq P/T \leq \frac{6\sqrt{\hat{\sigma}_{GRR}^2 + \sqrt{VU_{GRR}}}}{USL - LSL} \quad (3.60)$$

Signal-to-noise SNR ratio is a function of ρ_p . The MLS confidence interval for SNR is represented as

$$\sqrt{2L_{\rho_p}} \leq SNR \leq \sqrt{2U_{\rho_p}} \quad (3.61)$$

3.3.4 MLS confidence intervals for process capability metrics

Observed potential capability of the process C_p index is a function of σ_{Total}^2 . The MLS confidence interval for C_p is defined as

$$\frac{USL - LSL}{6\sqrt{\hat{\sigma}_{Total}^2 - \sqrt{VL_{Total}}}} \leq C_p \leq \frac{USL - LSL}{6\sqrt{\hat{\sigma}_{Total}^2 + \sqrt{VU_{Total}}}} \quad (3.62)$$

Actual potential capability of the process C_p^* is a function of σ_p^2 . The MLS confidence interval for C_p^* can be represented as

$$\frac{USL - LSL}{6\sqrt{\hat{\sigma}_p^2 - \sqrt{VL_p}}} \leq C_p^* \leq \frac{USL - LSL}{6\sqrt{\hat{\sigma}_p^2 + \sqrt{VU_p}}} \quad (3.63)$$

3.4 Balanced two-factor crossed random model without interaction

Significance of operator by part interaction effect can be determined using the F-test. Considering the following hypothesis testing:

$H_0 : \sigma_{po}^2 = 0$, implying that σ_{po}^2 is not significant

$H_1 : \sigma_{po}^2 > 0$, implying that σ_{po}^2 is significant

If $F_o = MS_{op}/MS_{Rep} < F - statistic = F_{\alpha, df_{op}, df_{Rep}}$ the null hypothesis cannot be rejected.

In this case the original model is modified to pool interaction with replication effect to fit a Two-factor crossed random model without interaction.

3.4.1 The model

In the balanced two-factor crossed random model with no interaction above, X_{ijk} represents the measurement done by appraiser i on sample j at replication k.

$$X_{ijk} = \mu + O_i + P_j + R_{k(ij)} \quad (3.64)$$

where $i = 1, \dots, o$; $j = 1, \dots, p$; $k = 1, \dots, n$.

In this model, μ is the overall mean of all observations, terms O_i and P_j represent the main effect of operator and part factors on the measurement respectively, and $R_{k(ij)}$ denotes the replication effect on X_{ijk} . Independency, randomness and normality behavior is assumed for these effects, where $O_i \sim N(0, \sigma_o^2)$, $P_j \sim N(0, \sigma_p^2)$ and $R_{k(ij)} \sim N(0, \sigma_{Rep}^2)$. In this particular model interaction effect is omitted, consequently mean square of repeatability is redefined as $MS_{Rep}^* = (SS_{op} + SS_{Rep})/df_{Rep}^*$, where $df_{Rep}^* = pon - p - o + 1$. ANOVA table for this model is shown on Table 3.5.

Table 3.5: ANOVA without interaction term

Source of variation	Sum of Squares	Degrees of Freedom	Mean Square	Expected mean square
P - Part	SS_p	$df_p = p - 1$	$MS_p = SS_p/df_p$	$\theta_p = \sigma_{Rep}^2 + on\sigma_p^2$
O - Operator	SS_o	$df_o = o - 1$	$MS_o = SS_o/df_o$	$\theta_o = \sigma_{Rep}^2 + pn\sigma_o^2$
R - Repeatability	$SS_{op} + SS_{Rep}$	$df_{Rep}^* = pon - p - o + 1$	$MS_{Rep}^* = (SS_{op} + SS_{Rep})/df_{Rep}^*$	$\theta_{Rep} = \sigma_{Rep}^2$
T - Total	SS_{Total}	$df_{Total} = pon - 1$		

Variance components are recalculated as follows

$$\hat{\sigma}_{Rep}^2 = MS_{Rep}^* \quad (3.65)$$

$$\hat{\sigma}_o^2 = \frac{MS_o - MS_{Rep}^*}{pn} \quad (3.66)$$

$$\hat{\sigma}_p^2 = \frac{MS_p - MS_{Rep}^*}{on} \quad (3.67)$$

As interaction effect between operator and part is not significant, reproducibility is equal to the variation due to operator effect:

$$\hat{\sigma}_{Reproducibility}^2 = \hat{\sigma}_o^2 \quad (3.68)$$

Measurement system variation is obtained as the sum of reproducibility and repeatability components

$$\hat{\sigma}_{GRR}^2 = \hat{\sigma}_{Rep}^2 + \hat{\sigma}_{Reproducibility}^2 = \frac{MS_o + (pn - 1)MS_{Rep}^*}{pn} \quad (3.69)$$

and total variation is the sum of all sources of variability in the system:

$$\hat{\sigma}_{Total}^2 = \hat{\sigma}_p^2 + \hat{\sigma}_{GRR}^2 = \frac{pMS_p + +oMS_o + (pon - p - o)MS_{Rep}^*}{pon} \quad (3.70)$$

3.4.2 MLS confidence intervals for variance components

Burdick and Larsen (1997) redefined the confidence interval of σ_{Rep}^2 when no interaction

effect is considered

$$(1 - G_{Rep}^*)MS_{Rep}^* \leq \sigma_{Rep}^2 \leq (1 + H_{Rep}^*)MS_{Rep}^* \quad (3.71)$$

Variability of operator σ_o^2 and manufacturing process σ_p^2 can be expressed as $\sigma_o^2 = (\theta_o - \theta_{Rep})/(pn)$ and $\sigma_p^2 = (\theta_p - \theta_{Rep})/(on)$. Ting et al. (1990) proposed a MLS method for calculating the CI of the difference of two expected mean squares. Applying this method, CI of σ_o^2 is expressed as

$$\hat{\sigma}_o^2 - \sqrt{VL_o} \leq \sigma_o^2 \leq \hat{\sigma}_o^2 + \sqrt{VU_o} \quad (3.72)$$

where

$$VL_o = \frac{G_o^2 MS_o^2 + (H_{Rep}^*)^2 (MS_{Rep}^*)^2 + G_{o,Rep}^* MS_o MS_{Rep}^*}{(pn)^2} \quad (3.73)$$

$$VU_o = \frac{H_o^2 MS_o^2 + (G_{Rep}^*)^2 (MS_{Rep}^*)^2 + H_{o,Rep}^* MS_o MS_{Rep}^*}{(pn)^2} \quad (3.74)$$

For the no interaction model $\sigma_{Reproducibility}^2 = \sigma_o^2$, hence the confidence interval shown above is the same for reproducibility. CI of σ_p^2 is

$$\hat{\sigma}_p^2 - \sqrt{VL_p} \leq \sigma_p^2 \leq \hat{\sigma}_p^2 + \sqrt{VU_p} \quad (3.75)$$

where

$$VL_p = \frac{G_p^2 MS_p^2 + (H_{Rep}^*)^2 (MS_{Rep}^*)^2 + G_{p,Rep}^* MS_p MS_{Rep}^*}{(on)^2} \quad (3.76)$$

$$VU_p = \frac{H_p^2 MS_p^2 + (G_{Rep}^*)^2 (MS_{Rep}^*)^2 + H_{p,Rep}^* MS_p MS_{Rep}^*}{(on)^2} \quad (3.77)$$

Total measurement system variability σ_{GRR}^2 and total variability σ_{Total}^2 can be calculated as $\sigma_{GRR}^2 = [\theta_o + (pn - 1)\theta_{Repeatability}]/(pn)$ and $\sigma_{Total}^2 = [p\theta_p + o\theta_o + (pon - p - o)\theta_{Repeatability}]/(pon)$ respectively. Graybill and Wang (1980) discussed a variation to the MLS method for calculating the addition of expected mean square values. Using this method CI of σ_{GRR}^2 is defined as

$$\hat{\sigma}_{GRR}^2 - \sqrt{VL_{GRR}} \leq \sigma_{GRR}^2 \leq \hat{\sigma}_{GRR}^2 + \sqrt{VU_{GRR}} \quad (3.78)$$

where

$$VL_{GRR} = \frac{G_o^2 MS_o^2 + (G_{Rep}^*)^2 (pn - 1)^2 (MS_{Rep}^*)^2}{(pn)^2} \quad (3.79)$$

$$VU_{GRR} = \frac{H_o^2 MS_o^2 + (H_{Rep}^*)^2 (pn - 1)^2 (MS_{Rep}^*)^2}{(pn)^2} \quad (3.80)$$

and CI of σ_{Total}^2 is calculated as

$$\hat{\sigma}_{Total}^2 - \sqrt{VL_{Total}} \leq \sigma_{Total}^2 \leq \hat{\sigma}_{Total}^2 + \sqrt{VU_{Total}} \quad (3.81)$$

where

$$VL_{Total} = [G_p^2 p^2 MS_p^2 + G_o^2 o^2 MS_o^2 + (G_{Rep}^*)^2 (pon - p - o)^2 (MS_{Rep}^*)^2]/(pon)^2 \quad (3.82)$$

$$VU_{Total} = [H_p^2 p^2 MS_p^2 + H_o^2 o^2 MS_o^2 + (H_{Rep}^*)^2 (pon - p - o)^2 (MS_{Rep}^*)^2] / (pon)^2 \quad (3.83)$$

3.4.3 MLS confidence intervals for measurement system capability metrics

Arteaga et al. (1982) proposed a modification of MLS method to calculate exact confidence intervals of the ratio of part variability to total measurement system variability ρ_p when Part-Operator interaction effect is not significant.

$$L_{\rho_p} \leq \rho_p \leq U_{\rho_p} \quad (3.84)$$

where

$$L_{\rho_p} = \frac{p(1 - G_p)MS_p^2 - pMS_pMS_{Rep}^* + p[I_{p,Rep}^* - (1 - G_p)(I_{p,Rep}^*)^2](MS_{Rep}^*)^2}{o(pn - 1)MS_pMS_{Rep}^* + o(1 - G_p)I_{p,o}MS_pMS_o} \quad (3.85)$$

$$U_{\rho_p} = \frac{p(1 - H_p)MS_p^2 - pMS_pMS_{Rep}^* + p[J_{p,Rep}^* - (1 + H_p)(J_{p,Rep}^*)^2](MS_{Rep}^*)^2}{o(pn - 1)MS_pMS_{Rep}^* + o(1 + H_p)J_{p,o}MS_pMS_o} \quad (3.86)$$

Burdick et al. (2005a) discussed a modification to MLS method to obtain exact confidence intervals of the ratio Repeatability to Total measurement system variability when interaction effect is omitted. CI of $\rho_{Repeatability}$ is presented as follows

$$\frac{L_{\rho_{Repeatability}}^*}{L_{\rho_{Repeatability}}^* + 1} \leq \rho_{Repeatability} \leq \frac{U_{\rho_{Repeatability}}^*}{U_{\rho_{Repeatability}}^* + 1} \quad (3.87)$$

where

$$L_{\rho_{Repeatability}}^* = \frac{1}{pn} \left(\frac{MS_o}{MS_{Rep}^* I_{o,Rep}^*} - 1 \right) \quad (3.88)$$

$$U_{\rho_{Repeatability}}^* = \frac{1}{pn} \left(\frac{MS_o}{MS_{Rep}^* J_{o,Rep}^*} - 1 \right) \quad (3.89)$$

Ratio $\rho_{Reproducibility} = 1 - \rho_{Repeatability}$. Considering previous result, the confidence interval of $\rho_{Reproducibility}$ can be defined as

$$1 - \frac{U_{\rho_{Repeatability}}^*}{U_{\rho_{Repeatability}}^* + 1} \leq \rho_{Reproducibility} \leq 1 - \frac{L_{\rho_{Repeatability}}^*}{L_{\rho_{Repeatability}}^* + 1} \quad (3.90)$$

Precision-to-tolerance P/T ratio is a function of total measurement system variability σ_{GRR}^2 ,

hence the MLS confidence interval for P/T is expressed as

$$\frac{6\sqrt{\hat{\sigma}_{GRR}^2 - \sqrt{V}L_{GRR}}}{USL - LSL} \leq P/T \leq \frac{6\sqrt{\hat{\sigma}_{GRR}^2 + \sqrt{V}U_{GRR}}}{USL - LSL} \quad (3.91)$$

Signal-to-noise SNR ratio is a function of ρ_p . The MLS confidence interval for SNR is calculated as

$$\sqrt{2L_{\rho_p}} \leq SNR \leq \sqrt{2U_{\rho_p}} \quad (3.92)$$

3.4.4 MLS confidence intervals for process capability metrics

Observed potential capability of the process C_p index is a function of σ_{Total}^2 . The MLS confidence interval for C_p can be expressed as

$$\frac{USL - LSL}{6\sqrt{\hat{\sigma}_{Total}^2 - \sqrt{VL_{Total}}}} \leq C_p \leq \frac{USL - LSL}{6\sqrt{\hat{\sigma}_{Total}^2 + \sqrt{VU_{Total}}}} \quad (3.93)$$

Actual potential capability of the process C_p^* is a function of σ_p^2 . The MLS confidence interval for C_p^* can be defined as

$$\frac{USL - LSL}{6\sqrt{\hat{\sigma}_p^2 - \sqrt{VL_p}}} \leq C_p^* \leq \frac{USL - LSL}{6\sqrt{\hat{\sigma}_p^2 + \sqrt{VU_p}}} \quad (3.94)$$

3.5 Algorithm

An algorithm based on the study performed by Pan (2004) was employed in order to calculate the point estimations and confidence intervals for the variance components and capability criteria. It is proposed to incorporate a test of significance for each of the sources of variation before estimating the variability components. Main goal of the calculations was to determine the effect of different levels of variability and allocation of resources in a classical Gauge R&R study on the width of each of the CIs estimated.

3.5.1 Assumptions and limitations

Some aspects of measurement system analysis, such as calibration and linearity assessing, are beyond the scope of the present research. Two scenarios are the base of this study: balanced two-factor crossed random model with interaction and balanced two-factor crossed random model with no interaction. Data for calculations are independently and randomly generated,

and it is assumed no shift of μ_T , and $\mu_{GRR} = 0$.

3.5.2 Procedure

The following procedure is established to calculate the Width for the confidence intervals of interest.

1. Set α level for Confidence interval calculations, mean and specification limits, i.e. $\alpha = 0.05$, $\mu_T = \mu_p = 150$, LSL= 100, USL= 200.

2. Establish the different combinations of sample size, number of operators and number of replications for the Gauge R&R study, i.e. $p \in \{5, 10, 15, 20, 25\}$, $o \in \{2, 3, 4, 5\}$ and $n \in \{2, 3, 4, 5\}$. As recommended by Antony et al. (1999) and He et al. (2000), these values represent the most common scenarios in industry.

3. Consider different levels for the variability components, i.e. $\sigma_{Rep}^2 \in \{1, 2, 3, 4, 5\}$, $\sigma_o^2 \in \{1, 2, 3, 4, 5\}$, $\sigma_{op}^2 \in \{1, 2, 3, 4, 5\}$ and $\sigma_p^2 \in \{1, 2, 3, 4, 5\}$. Following example from Pan (2004), sources of variation have the same values.

4. Calculate the expected mean square (EMS) values using the different combinations of factors established in step (2) and sources of variability in step (3). For example considering the combination $\{p = 15, o = 4, n = 3, \sigma_{Rep}^2 = 2, \sigma_o^2 = 3, \sigma_{op}^2 = 2, \sigma_p^2 = 5\}$, values of EMS are

$$\theta_{Rep} = \sigma_{Rep}^2 = 2$$

$$\theta_o = \sigma_{Rep}^2 + n\sigma_{op}^2 + pn\sigma_o^2 = 143$$

$$\theta_{op} = \sigma_{Rep}^2 + n\sigma_{op}^2 = 8$$

$$\theta_p = \sigma_{Rep}^2 + n\sigma_{op}^2 + on\sigma_p^2 = 68$$

5. Burdick and Larsen (1997) defined the ratio $df_q MS_q / \theta_q$ as an independent chi-square random variable with q degrees of freedom. For the present study the following ratios

are of interest: $df_{Rep}MS_{Rep}/\theta_{Rep} \sim \chi^2_{df_{Rep}}$, $df_oMS_o/\theta_o \sim \chi^2_{df_o}$, $df_{op}MS_{op}/\theta_{op} \sim \chi^2_{df_{op}}$ and $df_pMS_p/\theta_p \sim \chi^2_{df_p}$. Using data from step (2) the χ^2 random variables are generated and mean square values are calculated. For example, lets us assume that the following data is randomly generated $\chi^2_{df_{Rep}=120} = 118.64$, $\chi^2_{df_o=3} = 5.94$, $\chi^2_{df_{op}=42} = 40.34$ and $\chi^2_{df_p=14} = 13.57$. In this case mean square values are $MS_{Rep} = 1.98$, $MS_o = 283.28$, $MS_{op} = 7.68$ and $MS_p = 65.92$.

6. Mean square values from step (5) and degrees of freedom calculated with data from step (2) are used to perform the Hypothesis testing in order to validate significance of the sources of variation. Considering part-operator interaction effect, if $F_o = MS_{op}/MS_{Rep} < F - statistic = F_{\alpha, df_{op}, df_{Rep}}$ the null hypothesis cannot be rejected and consequently the balanced two-factor crossed random model with no interaction is considered for calculating the following steps, otherwise the balanced two-factor crossed random model with interaction is established for calculating the variance and capability criteria point estimations and their respective Confidence intervals. In the present example $(MS_{op}/MS_{Rep}) = 3.88 > F_{\alpha=0.05, df_{op}=42, df_{Rep}=120} = 1.49$, implying that part-operator interaction effect is indeed significant.

7. Calculate the point estimations of the variability components $\hat{\sigma}_{Rep}^2$, $\hat{\sigma}_o^2$, $\hat{\sigma}_{op}^2$, $\hat{\sigma}_p^2$, $\hat{\sigma}_{Reproducibility}^2$, $\hat{\sigma}_{GRR}^2$, $\hat{\sigma}_{Total}^2$ and capability criteria $\hat{\rho}_p$, $\hat{\rho}_{Rep}$, $\hat{\rho}_{Reproducibility}$, $\widehat{P/T}$, \widehat{SNR} , \hat{C}_p , \hat{C}_p^* .
8. Determine the confidence intervals for the variability components and capability criteria from step (7).
9. Repeat steps (5) to (8) 100.000 times. Use the lower and upper bound averages to calculate the width of the confidence interval for each variability component and capability criterion. For a given point estimation ω , with an upper bound U and lower bound L, the width is

determined with the following equation

$$Width = \frac{U - L}{\omega} \quad (3.95)$$

A total of 50.000 different combinations of variability components and resources for a Gauge R&R study were simulated using the algorithm described above. Analysis of results obtained are presented in next chapter.

Chapter 4

Sensitivity analysis of variability components and capability criteria

The main purpose of this research is to study using numerical simulation the average confidence interval width of variability components and capability metrics under different parameter allocations and various levels of variability. The simulation study considered a total of 50,000 scenarios. In order to ensure reliability of the results, each average CI width was calculated from a total of 100,000 independent replications. An example of the data obtained from the simulation study is shown in *Appendix A*.

Present chapter offers a discussion about the most important findings from a selected group of cases and provides guidance for designing a more advantageous Gauge R&R study depending on the metrics of interest for practitioners. Findings are then compared with a real case of study presented in the next chapter.

Table 4.1 shows the 80 different allocations of parameters, sample size, number of operators and number of replications included in the sensitivity analysis. The range of values for each parameter is considered to be the most commonly used in industry, making conclusions from the present research an useful guidance for practical environment. Total measurement number differs from a minimum of 20 considering a parameter allocation combination of $(p=5, o=2, n=2)$, to a maximum of 625 with the parameter allocation combination $(p=25, o=5, n=5)$. In addition to this, each design was tested with respect to 625 different

Table 4.1: Total measurement number for different parameter allocations

Sample size (p)	Number of operators (o)	Replications (n)			
		2	3	4	5
5	2	20	30	40	50
5	3	30	45	60	75
5	4	40	60	80	100
5	5	50	75	100	125
10	2	40	60	80	100
10	3	60	90	120	150
10	4	80	120	160	200
10	5	100	150	200	250
15	2	60	90	120	150
15	3	90	135	180	225
15	4	120	180	240	300
15	5	150	225	300	375
20	2	80	120	160	200
20	3	120	180	240	300
20	4	160	240	320	400
20	5	200	300	400	500
25	2	100	150	200	250
25	3	150	225	300	375
25	4	200	300	400	500
25	5	250	375	500	625

variability component combinations.

4.1 Analysis of variance components

Gauge R&R study focuses on interpreting all variability components in the manufacturing and measurement system. Point estimators and confidence intervals for all of these values can be calculated using ANOVA method. Depending on the design of the Gauge R&R study, the width of the confidence intervals can greatly change. Therefore, results from a poor design might not be of use for concluding about reliability of the system.

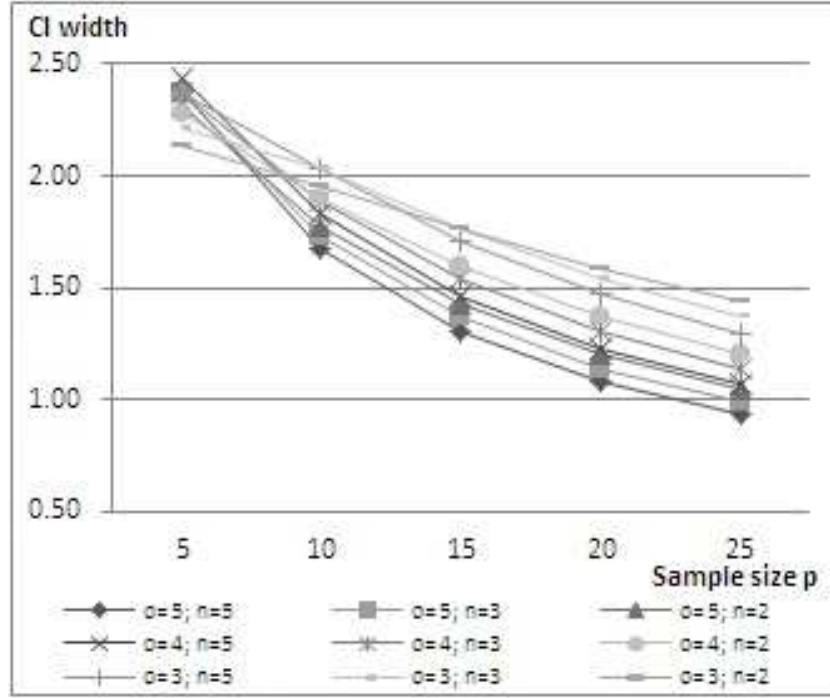
This section gives understanding of how the common parameters used in a Gauge R&R study, such as sample size, number of operators and number of replications, can influence the width of the confidence interval of each variability component and capability criteria. The research will allow the practitioner to have better understanding for choosing a proper parameter allocation design based on the metric of interest.

4.1.1 Analysis of repeatability σ_{Rep}^2

From Figure 4.1, it is clear that the width of the confidence interval of repeatability σ_{Rep}^2 , the first component of the Gauge R&R study, presents a significant reduction when increasing sample size in the design. To illustrate this, considering the design with sample size $p=10$, number of operators $o=3$ and replications $n=3$, the resulted average confidence interval width is 2.03. By increasing the sample size to 15, the average confidence interval width of the design ($p=15, o=3, n=3$) changes to 1.76. This new value represents a reduction of 13.2% in comparison to the original average width.

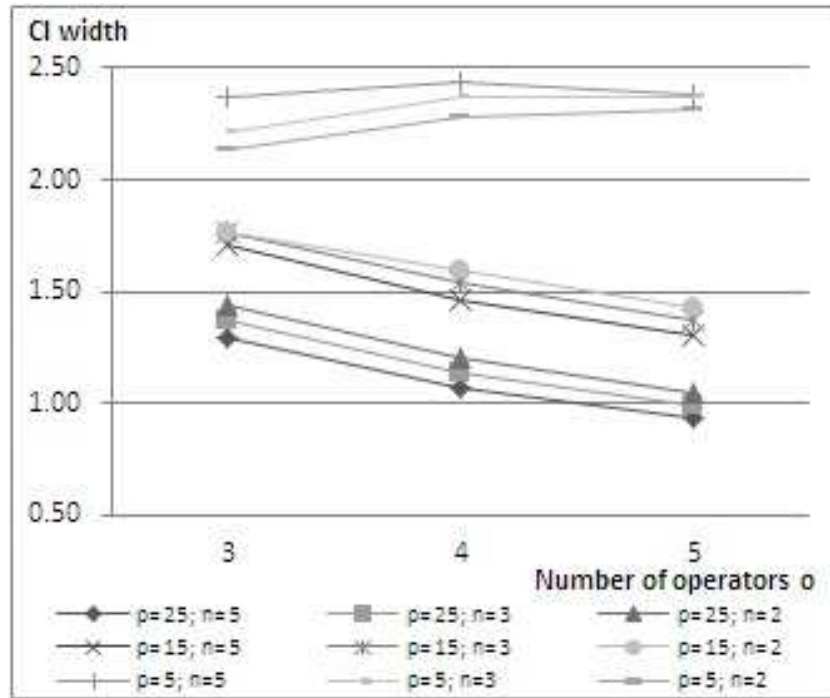
Figure 4.2 presents the effect of number of operators on the average confidence interval width. Twenty four out of twenty seven parameter allocation combinations show a moderate downward trend when increasing the number of operators in the design. For example, the design ($p=15, o=4, n=2$) has a average CI width of 1.59. By adding 1 operator to previous design, ($p=15, o=5, n=2$), the average CI width changes to 1.43, implying a reduction of 10.1%. The combination of the lowest sample size and smallest number of operators, ($p=5, o=3, n=5$), ($p=5, o=3, n=3$) and ($p=5, o=3, n=2$), presents an atypical upward trend when increasing the number of operators in the experiment.

Figure 4.3 also shows some atypical upward trend when increasing the number of replications



* Average values from 625 different combinations of variability components

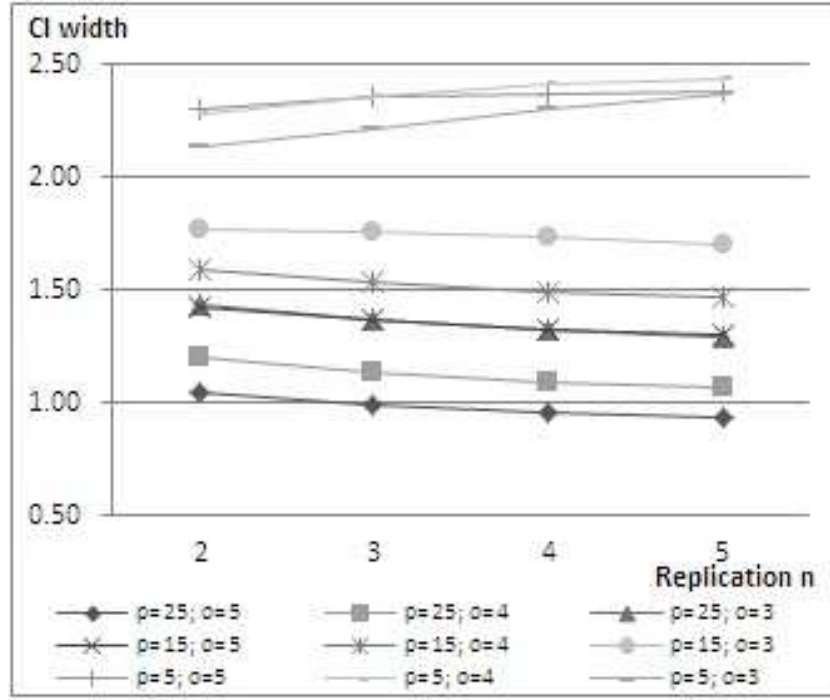
Figure 4.1: Average CI width of σ^2_{Rep} for different sample size p



* Average values from 625 different combinations of variability components

Figure 4.2: Average CI width of σ^2_{Rep} for different sample size p

in the design. This occurrence is noticed in all scenarios with the smallest sample size. Most of the results indicate that there is a small change on the average confidence interval width when changing the number of replications. As an example, the design ($p=25, o=5, n=3$) returns a average CI width of 0.99. And by increasing 1 replication in previous design, ($p=25, o=5, n=4$), it is obtained an average width of 0.96, that is 3.1% narrower than the original average width.



* Average values from 625 different combinations of variability components

Figure 4.3: Average CI width of σ_{Rep}^2 for different replication n

Table 4.2 shows the average confidence interval width calculated for different parameter allocations and various levels of variability components. To illustrate this analysis, designs were separated in three groups with 120, 200 and 300 total measurements. Conclusions drawn from this analysis can be considered as a reference for deciding the best parameter allocation combination when the total number of experiments is limited, most commonly because of cost or time restrictions.

Table 4.2 indicates that when having a limited number of measurements in the Gauge R&R study, it is recommended to select parameter allocation combinations with larger number of replications. To illustrate, considering the case with 300 measurements and variability components $\sigma_{Rep}^2 = 5$, $\sigma_p^2 = 5$, $\sigma_o^2 = 3$ and $\sigma_{op}^2 = 3$, narrower confidence intervals are obtained with the maximum number of replications, (p=15,o=4,n=5) and (p=20,o=3,n=5). For both scenarios the average CI width is 0.36, that represents a reduction of 7.7% with respect to the larger average confidence interval width.

Table 4.2: Average CI width of σ_{Rep}^2 for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	0.61	0.61	0.61	0.59	0.61	0.61	0.57	0.60	0.61
	10	4	3	0.65	0.65	0.65	0.62	0.65	0.65	0.58	0.64	0.65
	15	4	2	0.76	0.76	0.76	0.66	0.76	0.76	0.61	0.73	0.76
	20	3	2	0.76	0.76	0.76	0.66	0.76	0.76	0.61	0.73	0.76
200	10	4	5	0.45	0.45	0.45	0.45	0.45	0.45	0.44	0.45	0.45
	10	5	4	0.46	0.46	0.46	0.46	0.46	0.46	0.45	0.46	0.46
	20	5	2	0.58	0.58	0.58	0.54	0.58	0.58	0.49	0.57	0.58
	25	4	2	0.58	0.58	0.58	0.54	0.58	0.58	0.49	0.57	0.58
300	15	4	5	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36
	15	5	4	0.38	0.38	0.38	0.38	0.38	0.38	0.37	0.38	0.38
	20	3	5	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36
	20	5	3	0.40	0.40	0.40	0.40	0.40	0.40	0.39	0.40	0.40
	25	3	4	0.38	0.38	0.38	0.37	0.38	0.38	0.37	0.38	0.38
	25	4	3	0.40	0.40	0.40	0.40	0.40	0.40	0.38	0.40	0.40

4.1.2 Analysis of part variability σ_p^2

Confidence interval of σ_p^2 is highly affected by df_p . Figure 4.4 shows that increasing the number of samples in the Gauge R&R study reduces dramatically the width of the confidence interval. For example, considering the scenario (p=5,o=5,n=5), the resulting average CI width obtained is 9.51. By adding 5 samples, the new design (p=10,o=5,n=5) reduces the

average CI width to 3.73. This signifies a reduction of 60.8% from the original design. It also can be noticed that the reduction pattern of average CI width slows down when sample size becomes larger, e.g., when $p=20, 25$.

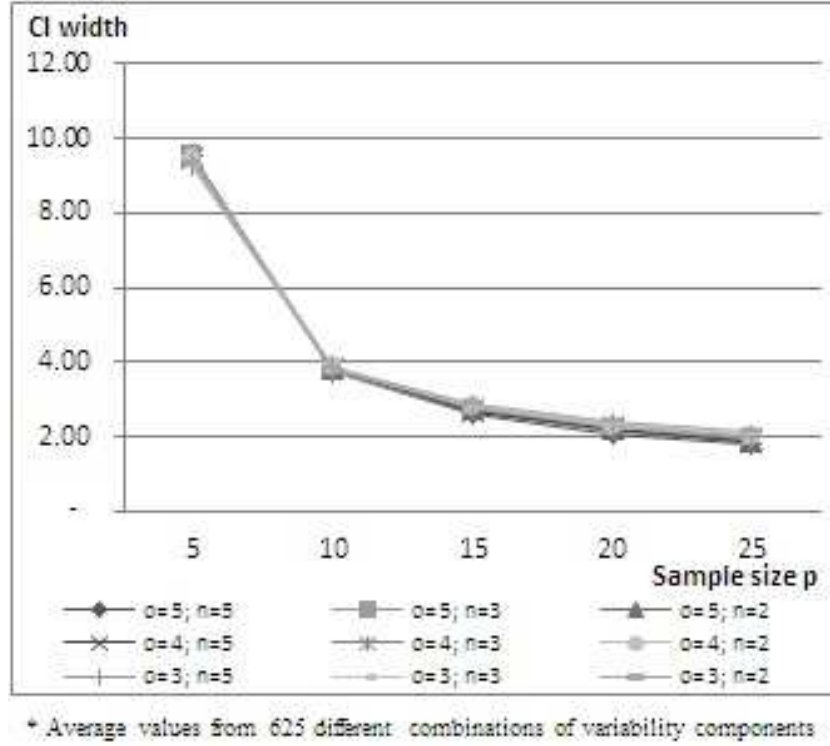
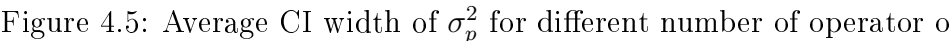


Figure 4.4: Average CI width of σ_p^2 for different sample size p

Figure 4.5 offers a clear perspective on the influence of the number of operators over the reduction of the average CI width. It is shown that an increase in the number of operators does not reduce significantly the average confidence interval width. Taking as an example the parameter allocation combination ($p=15, o=3, n=5$), the resulting average CI width is 2.80. By including 1 more operator, the average width of the new design ($p=15, o=4, n=5$) changes to 2.71, representing a reduction of only 3.2%. It can also be observed that for scenarios with the same number of operators, increasing number of parts has to be prioritized instead of number of replications for obtaining a better outcome.

Results in Figure 4.6 demonstrate that the number of replications has virtually no effect



on the reduction of the average CI width. As an example, design (p=25,o=5,n=4) returns an average width of 1.83, while design (p=25,o=5,n=5) returns an average width of 1.81. Increasing 1 replication in the design reduces the width of the confidence interval in just 1.0%. It is also evident from the graphic that when considering designs with the same number of replications, a change on the numbers of parts affects more the width of the confidence interval than a change on the number of operators.

Table 4.3 shows the behavior of the average CI width for total measurement numbers equal to 120, 200 and 300. It is interesting to notice that, for a fixed number of total measurements, a larger sample size is always of benefit for reducing the width of CI. As an example, considering $\sigma_{Rep}^2 = 1$, $\sigma_p^2 = 5$, $\sigma_o^2 = 1$ and $\sigma_{op}^2 = 3$, the narrower confidence interval for 200 measurements is obtained with the parameter allocation combination (p=25,o=4,n=2) that has an average width of 1.60, less than half the value from the worst design (p=10,o=4,n=5) that presents

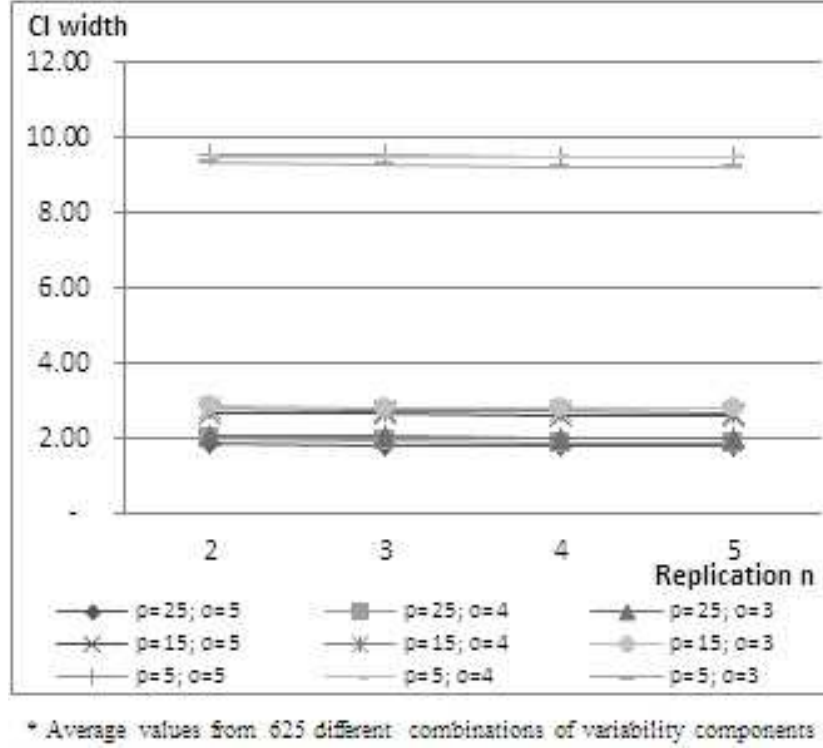


Figure 4.6: Average CI width of σ_p^2 for different replications n

an average width of 3.44. All combinations with smaller sample sizes present wider CIs. These results confirm that when part variability is considered fundamental in the Gauge R&R study, the experiment has to be designed in a way to include a considerable amount of samples.

4.1.3 Analysis of operator variability σ_o^2

With respect to operator variability σ_o^2 , the width of its confidence interval presents a small reduction with respect to an increment in the sample size. As an illustration, from Figure 4.7 design (p=20,o=3,n=5) has a average CI width of 43.38. By adding 5 samples, , the average width of the new design (p=25,o=3,n=5) changes to 43.02. This is a reduction of just 0.8%. From the graphic it can be also concluded that for a fix number of samples, it is preferable to increase the number of operators rather than the number of replications to reduce the width

Table 4.3: Average CI width of σ_p^2 for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	3.88	3.56	4.01	3.92	3.66	4.05	3.90	3.70	4.07
	10	4	3	3.86	3.46	4.01	3.98	3.58	4.07	4.01	3.65	4.11
	15	4	2	2.80	2.38	2.91	2.98	2.51	3.01	3.07	2.60	3.08
	20	3	2	2.42	2.01	2.53	2.53	2.14	2.61	2.62	2.23	2.66
200	10	4	5	3.81	3.44	3.99	3.95	3.53	4.04	3.97	3.58	4.06
	10	5	4	3.73	3.35	3.92	3.92	3.44	3.98	3.98	3.50	4.02
	20	5	2	2.13	1.81	2.26	2.42	1.90	2.36	2.55	1.98	2.45
	25	4	2	1.94	1.60	2.06	2.22	1.69	2.17	2.33	1.77	2.26
300	15	4	5	2.67	2.34	2.88	2.84	2.40	2.93	2.94	2.45	2.97
	15	5	4	2.57	2.27	2.77	2.77	2.32	2.83	2.90	2.37	2.88
	20	3	5	2.30	1.96	2.50	2.46	2.03	2.54	2.53	2.08	2.57
	20	5	3	2.07	1.80	2.23	2.31	1.85	2.31	2.46	1.91	2.38
	25	3	4	2.01	1.67	2.20	2.21	1.74	2.26	2.29	1.80	2.30
	25	4	3	1.87	1.58	2.04	2.12	1.64	2.12	2.25	1.70	2.19

of the confidence interval.

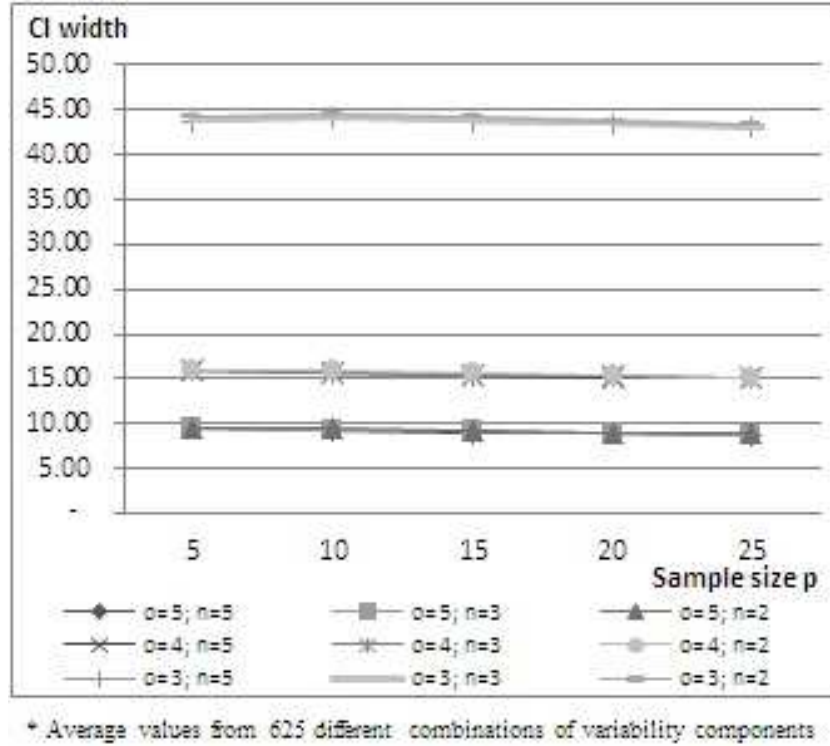


Figure 4.7: Average CI width of σ_o^2 for different sample size p

From Figure 4.8 it can be noticed that the average CI width of σ_o^2 is mainly affected by df_o .

Regardless the sample size or number of replications in the design, augmenting the number of operators highly reduces the average width. Considering parameter allocation combination ($p=15, o=4, n=2$), the average width obtained is 15.61. Modifying the number of operators in the design to ($p=15, o=5, n=2$) returns a new average width of 9.18, that results in a reduction of 41.2% with respect to the original average CI width.

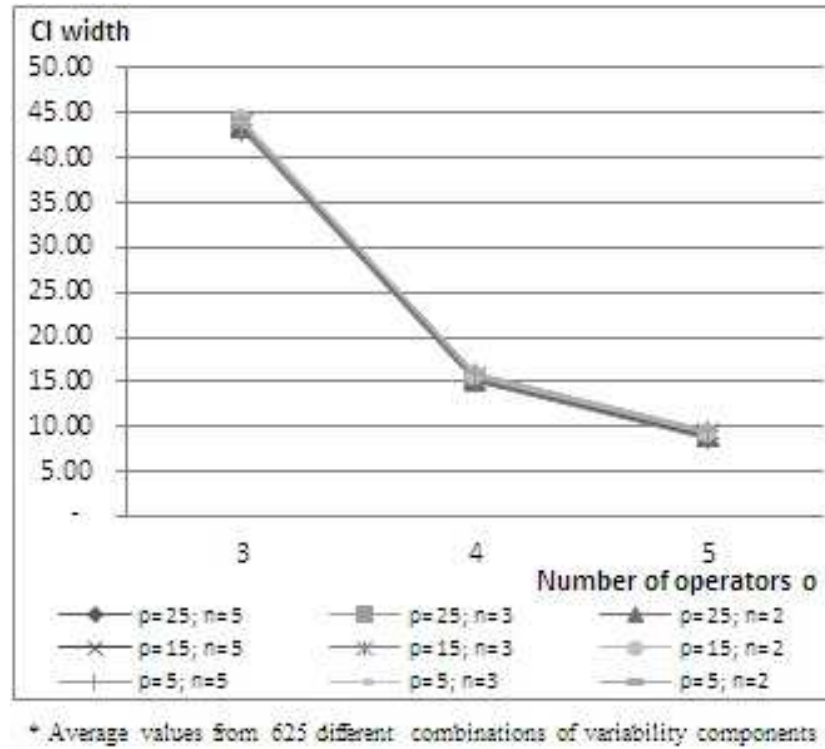
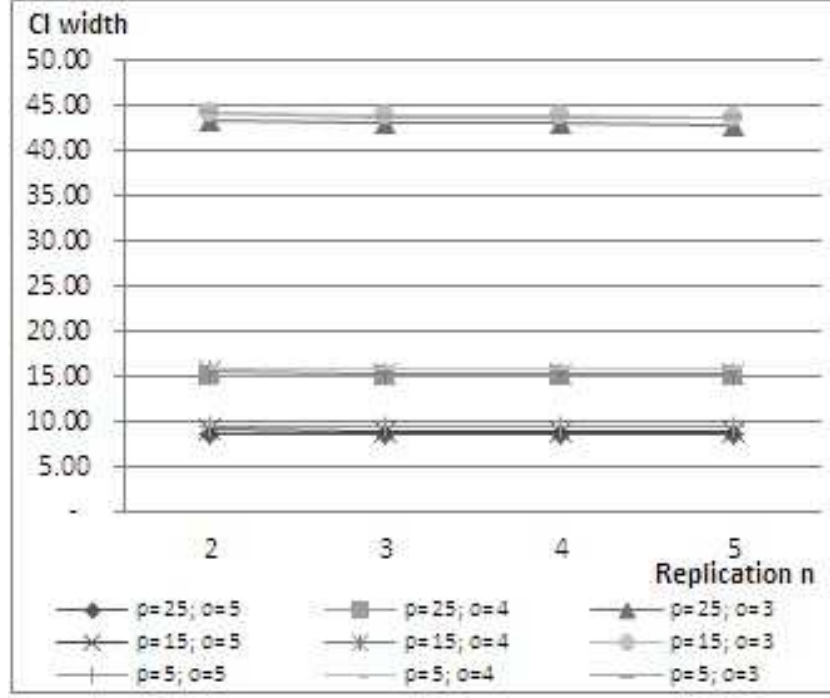


Figure 4.8: Average CI width of σ_o^2 for different number of operator o

It is observed in Figure 4.9 that increasing the number of replications does not significantly alter the width of the confidence interval. Considering the design ($p=5, o=3, n=4$), the resulting average CI width is 43.74. And for the design ($p=5, o=3, n=5$) the average width changes to 43.73. By augmenting 1 replication to the design, that implies passing from 60 to 75 measurements in the Gauge R&R study, the reduction on the average width is 0.0%. It is evident in this graphic that when the number of replications are fixed, the best decision is to increase the number of operators instead of the sample size.



* Average values from 625 different combinations of variability components

Figure 4.9: Average CI width of σ_o^2 for different replication n

Table 4.4 presents the average CI width for fixed number of total measurements and various levels of variability components. Results show that designs that maximize number of operators despite other parameters tend to have the narrowest confidence intervals width. Considering as an example a total number of 300 measurements and variability components $\sigma_{Rep}^2 = 3$, $\sigma_p^2 = 1$, $\sigma_o^2 = 3$ and $\sigma_{op}^2 = 1$, designs (p=15,o=5,n=4) with average CI width of 8.46 and design (p=20,o=5,n=3) with average CI width of 8.39 are the best options. Last design represents a reduction of 79.9% with respect to the average width of the worst parameter allocation combination.

It is important to emphasize that in the present research, a total of 12,500 different scenarios with number of operators o=2 were simulated. The average CI widths obtained were excessively large in all cases, and variability components and capability metrics that are functions of operator variability presented similar results. This indicates that when using MLS method

Table 4.4: Average CI width of σ_o^2 for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	44.23	46.5	44.92	42.37	43.31	43.95	44.79	44.56	44.08
	10	4	3	15.76	16.84	16.07	14.92	15.24	15.59	16.22	15.94	15.70
	15	4	2	15.53	16.60	15.76	14.81	15.01	15.35	16.31	15.73	15.49
	20	3	2	43.63	45.82	44.07	42.10	42.60	43.24	45.12	44.05	43.50
200	10	4	5	15.65	16.79	16.04	14.80	15.13	15.54	16.02	15.79	15.61
	10	5	4	9.24	10.09	9.53	8.66	8.85	9.16	9.61	9.36	9.22
	20	5	2	8.87	9.62	9.03	8.48	8.55	8.77	9.50	9.04	8.87
	25	4	2	15.03	16.03	15.24	14.48	14.58	14.89	15.83	15.24	15.02
300	15	4	5	15.29	16.47	15.68	14.54	14.80	15.20	15.82	15.43	15.28
	15	5	4	8.95	9.80	9.22	8.46	8.60	8.88	9.41	9.05	8.94
	20	3	5	43.17	45.56	43.96	41.65	42.22	42.97	44.13	43.43	43.13
	20	5	3	8.79	9.58	9.01	8.39	8.48	8.72	9.31	8.91	8.79
	25	3	4	42.85	45.15	43.53	41.50	41.90	42.65	43.96	43.14	42.81
	25	4	3	14.92	15.96	15.20	14.36	14.50	14.82	15.59	15.08	14.91

for calculating confidence intervals, the design of the Gauge R&R study must consider a minimum of 3 operators for obtaining meaningful results. When in practice it is not possible to consider more than 2 operators in the design, practitioner must refer to other methods such as Satterthwaite approximation.

4.1.4 Analysis of operator by part variability σ_{op}^2

The width of the confidence interval of operator by part variability is influenced especially by df_p and df_o in the Gauge R&R study. From Figure 4.10, it can be interpreted that increasing the sample size has a remarkable impact on the reduction of the confidence interval width. Previous conclusion can be demonstrated considering the parameter allocation combination (p=10,o=3,n=5) that gives an average width of 2.19. Changing the number of samples in previous design to (p=15,o=3,n=5) modifies the average width to 1.60. With five more samples in the Gauge R&R study, a reduction of 26.9% in the confidence interval is obtained.

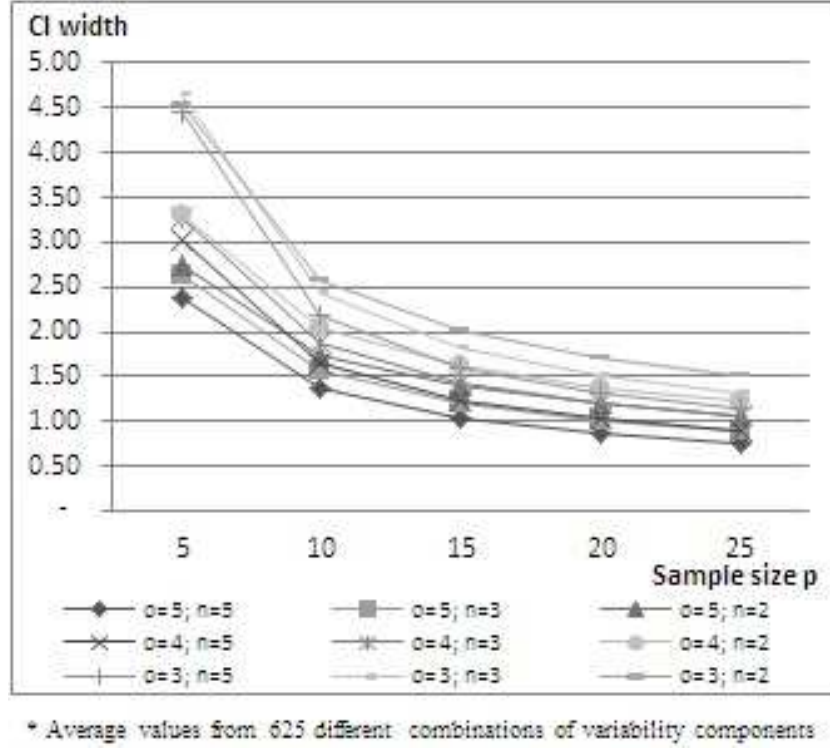


Figure 4.10: Average CI width of σ_{op}^2 for different sample size p

Figure 4.11 corroborates that there exist a significant influence of number of operators on the width of the confidence interval. To illustrate this finding, it can be shown that parameter allocation combination (p=15,o=3,n=2) produces an average CI width of 2.01. By increasing 1 operator the average width of the new design (p=15,o=4,n=2) is reduced to 1.61, implying a decrement of 19.8%. It is also important to indicate that when having the same number of operators, scenarios with the smallest sample size, i.e. p=5, have considerably larger confidence intervals than the other scenarios.

Number of replications seems to moderately affect the confidence interval width as seen in Figure 4.12. As an example, parameter allocation combination (p=25,o=4,n=2) presents an average CI width of 1.22. By augmenting 1 replication, the average width of the new design (p=25,o=4,n=3) reduces to 1.05. This represents a reduction of 14.3% with respect to the previous confidence interval. The graphic also shows that scenarios with sample size p=5 are

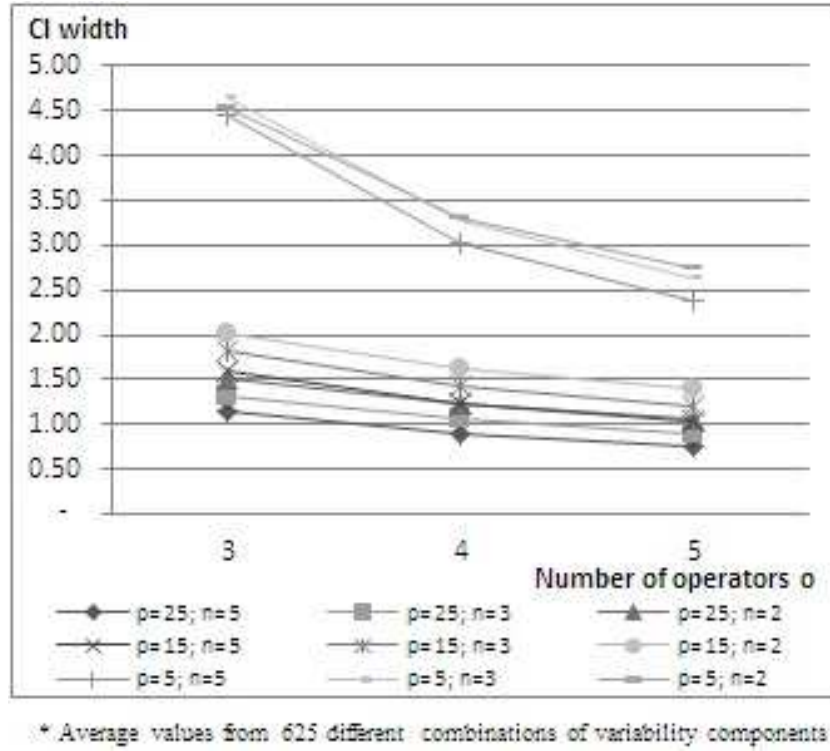


Figure 4.11: Average CI width of σ_{op}^2 for different number of operator o

the ones with the wider confidence intervals.

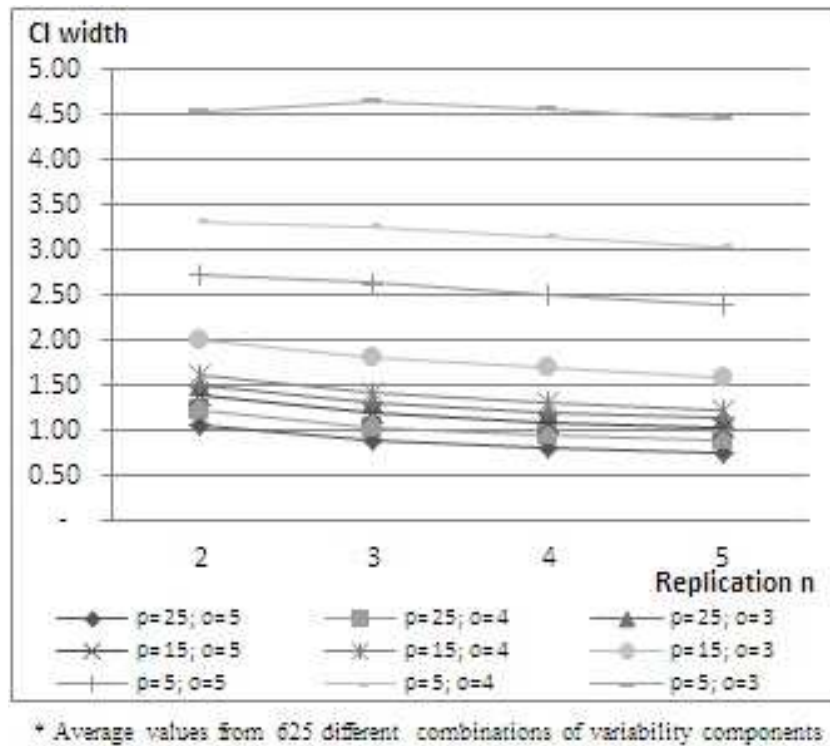


Figure 4.12: Average CI width of σ_{op}^2 for different replication n

Considering a fixed number of total measurements, narrower confidence intervals are obtained when the design of the Gauge R&R study reduces the number of replications to increase sample size or number of operators. From Table 4.5, for the scenario with 200 total measurement number and $\sigma_{Rep}^2 = 3$, $\sigma_p^2 = 3$, $\sigma_o^2 = 5$ and $\sigma_{op}^2 = 5$, combination (p=20,o=5,n=2) returns an average width of 0.90 and combination (p=25,o=4,n=2) an average width of 0.93. For designs with larger number of replication the confidence intervals get wider. Scenario (p=10,o=4,n=5) gives an average width of 1.41 and combination (p=10,o=5,n=4) one of 1.20.

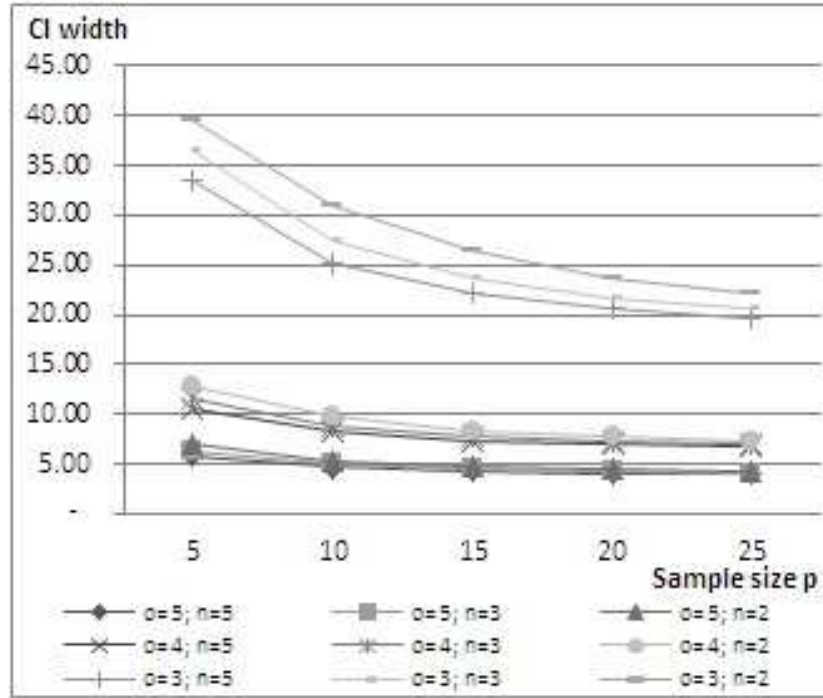
Table 4.5: Average CI width of σ_{op}^2 for different total measurement numbers

Measurements	p	o	σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
			σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	2.21	1.79	1.73	2.97	2.16	1.98	3.25	2.52	2.26
	10	4	3	1.77	1.39	1.33	2.47	1.75	1.55	2.7	2.09	1.8
	15	4	2	1.56	1.12	1.04	2.16	1.55	1.3	2.31	1.87	1.59
	20	3	2	1.67	1.19	1.11	2.27	1.64	1.4	2.41	1.97	1.71
200	10	4	5	1.53	1.32	1.29	2.16	1.52	1.41	2.55	1.74	1.54
	10	5	4	1.33	1.12	1.08	1.96	1.32	1.2	2.32	1.55	1.34
	20	5	2	1.09	0.79	0.74	1.79	1.09	0.9	1.99	1.4	1.09
	25	4	2	1.12	0.82	0.76	1.82	1.12	0.93	2.03	1.44	1.13
300	15	4	5	1.15	1	0.98	1.62	1.14	1.06	2.04	1.29	1.15
	15	5	4	1.01	0.86	0.83	1.51	1.01	0.92	1.91	1.17	1.01
	20	3	5	1.22	1.07	1.04	1.74	1.21	1.13	2.16	1.38	1.23
	20	5	3	0.92	0.75	0.71	1.49	0.92	0.81	1.86	1.1	0.92
	25	3	4	1.11	0.94	0.91	1.67	1.11	1.01	2.08	1.28	1.12
	25	4	3	0.94	0.77	0.74	1.53	0.94	0.84	1.91	1.14	0.95

These results indicate that when designing a Gauge R&R study, priority to the sample size and number of operators has to be given instead to number of replications when operator by part variability is considered to be significant for the experiment.

4.1.5 Analysis of reproducibility $\sigma_{Reproducibility}^2$

Reproducibility $\sigma_{Reproducibility}^2$, second component of Gauge R&R study, is composed of σ_o^2 and σ_{op}^2 . It can be verified in the following graphics that the CI of $\sigma_{Reproducibility}^2$ depends on df_o and df_p . Figure 4.13 demonstrates that increasing the sample size helps to a small reduction on the average CI width, but it is more evident when having the lowest number of operators. Considering the parameter allocation combination (p=20,o=4,n=3) it is obtained an average width of 7.21, and for design (p=25,o=4,n=3) a value of 6.92. This represents a reduction of 4.1%. It is clearly evidenced that all scenarios with just 3 operators have average CI widths more than twice the width of CI with larger number of operators.



* Average values from 625 different combinations of variability components

Figure 4.13: Average CI width of $\sigma_{Reproducibility}^2$ for different sample size p

Regardless the combination of number of replications and sample size, having more operators in the design highly reduces the width of the confidence interval. From Figure 4.14, it is shown

that combination (p=25,o=3,n=2) presents an average of 22.16, and by increasing 1 operator, the new design (p=25,o=4,n=2) obtains an average width of 7.30. This signifies a reduction of 67.0%.

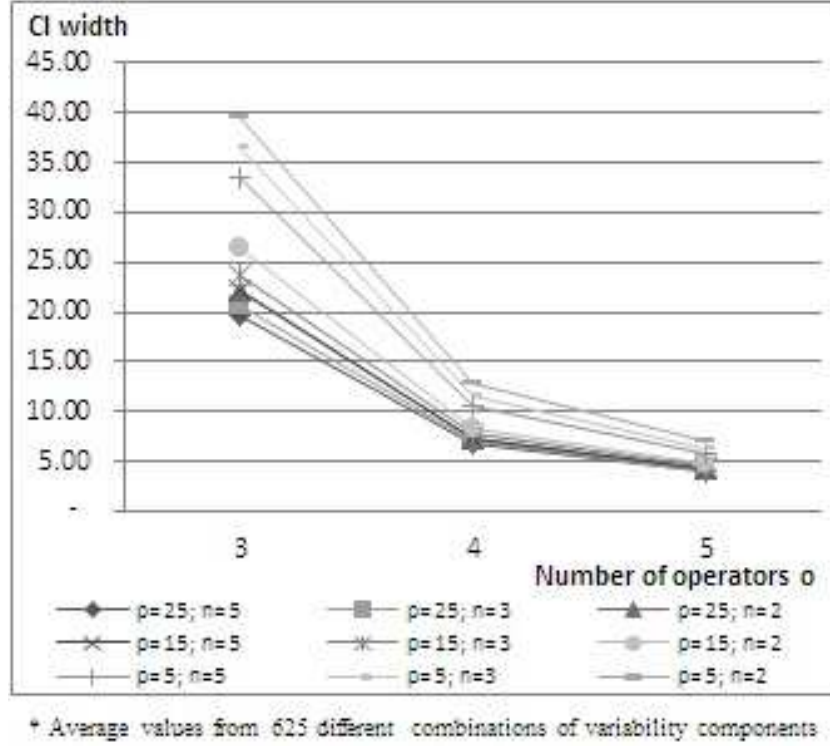


Figure 4.14: Average CI width of $\sigma_{Reproducibility}^2$ for different number of operator o

An increase in the number of replications causes a slight reduction on width of the CI. It can be illustrated with Figure 4.15. Parameter allocation combination (p=15,o=3,n=3) has an average CI width of 23.70. In the new design (p=15,o=3,n=4) it changes to 22.71, obtaining a reduction of just 4.2%. When the number of operators in the design is small, i.e. o=3, the average CI width presents a larger increase.

When reproducibility is fundamental for the analysis of the Gauge R&R study, number of operators has to be prioritized when designing the experiment. In Table 4.6, scenario where the total number of measurements is 200 and the components of variability are $\sigma_{Rep}^2 = \sigma_p^2 = \sigma_o^2 = \sigma_{op}^2 = 1$, best designs are the ones with the maximum number of operators. Combination

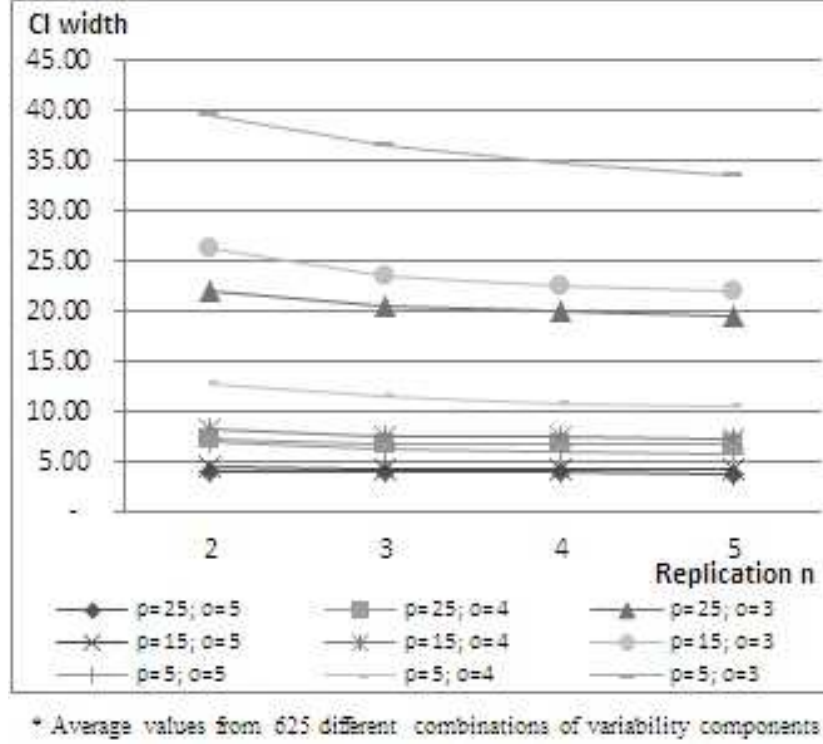


Figure 4.15: Average CI width of $\sigma^2_{Reproducibility}$ for different replication n

(p=10,o=5,n=4) and (p=20,o=5,n=2) have an average width of 4.53 and 4.13 respectively. Last combination decreases the width of the confidence interval 47.7% with respect to the worst design. It is also noticed that when number of operators cannot be modified, then increase of sample size has to be considered instead of number of replications as it also influences the CI width reduction.

4.1.6 Analysis of total measurement system variability σ^2_{GRR}

Total measurement system variability σ^2_{GRR} results from the summation of repeatability and reproducibility. It is considered the principal error estimation to analyze in a Gauge R&R study. Results from Figure 4.16 show that sample size has a moderate effect on the reduction of the CI width. For example, considering design (p=15,o=4,n=2) with average width equal to 5.12. Augmenting 5 parts, the average width of the combination (p=20,o=4,n=2) is

Table 4.6: Average CI width of $\sigma_{Reproducibility}^2$ for different total measurement numbers

Measurements	p	o	σ_{op}^2 n	σ_{Rep}^2	1	1	1	3	3	3	5	5	5
				σ_p^2	1	5	3	1	5	3	1	5	3
				σ_o^2	1	1	3	3	5	5	1	3	5
				n	1	3	5	1	3	5	1	3	5
120	10	3	4	4	24.82	19.01	21.20	34.00	26.41	24.33	39.25	26.23	25.35
	10	4	3	3	8.21	6.07	6.81	11.69	8.98	8.01	13.84	8.84	8.36
	15	4	2	2	7.68	5.16	6.02	12.06	8.59	7.38	14.35	8.73	7.82
	20	3	2	2	21.66	14.38	17.01	35.07	24.10	20.85	40.85	24.73	22.25
200	10	4	5	5	7.90	5.94	6.74	10.54	8.75	7.80	10.93	8.12	7.99
	10	5	4	4	4.53	3.32	3.78	6.17	5.10	4.46	6.49	4.70	4.58
	20	5	2	2	4.13	2.61	3.21	6.31	4.82	4.02	7.24	4.38	4.15
	25	4	2	2	6.83	4.29	5.33	10.66	8.02	6.68	12.22	7.27	6.90
300	15	4	5	5	7.14	4.97	5.93	9.86	8.22	7.08	8.89	7.28	7.19
	15	5	4	4	4.15	2.79	3.37	5.84	4.84	4.10	5.31	4.27	4.17
	20	3	5	5	20.10	13.86	16.70	27.60	22.92	19.94	25.61	20.46	20.27
	20	5	3	3	4.02	2.56	3.18	5.80	4.76	3.96	5.47	4.14	4.04
	25	3	4	4	19.36	12.86	15.87	27.22	22.46	19.20	25.82	19.79	19.58
	25	4	3	3	6.68	4.22	5.30	9.69	7.89	6.59	9.24	6.86	6.72

modified to 4.81. This is a reduction of 6.1% from the original value.

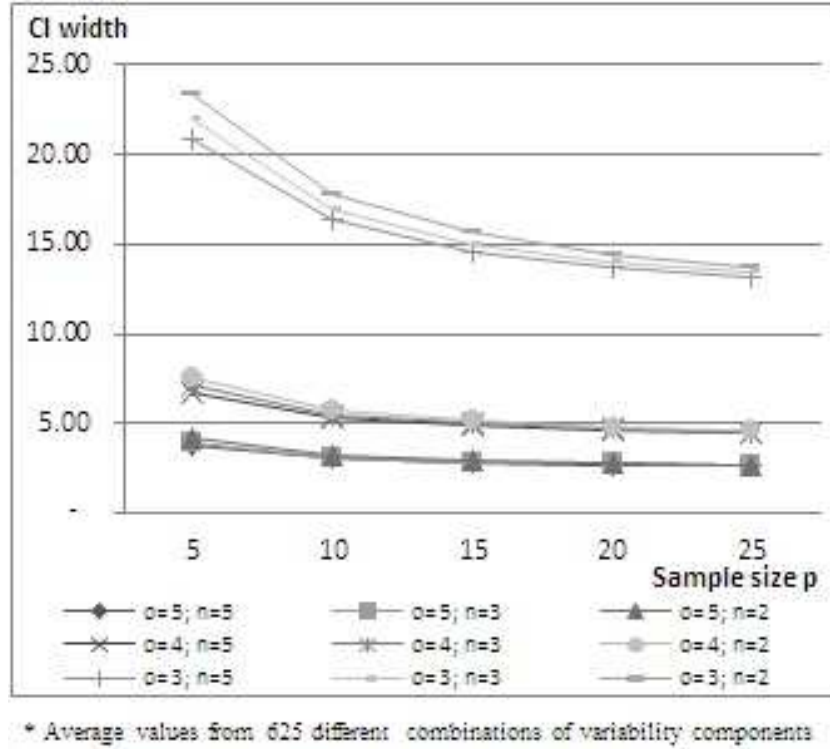


Figure 4.16: Average CI width of σ_{GRR}^2 for different sample size p

Most important parameter to analyze is the number of operators. Figure 4.17 clearly demon-

states that increasing this parameter has a high effect on reducing the width of the confidence interval. This is evident for the design ($p=25, o=4, n=3$) that results on an average width of 4.53. Adding 1 operator, the new design ($p=25, o=5, n=3$) reduces its average width to 2.64, which is an improvement of 41.7% compared to the original CI width. The graphic also shows that for the same number of operators, designs with smaller sample size have wider CI.

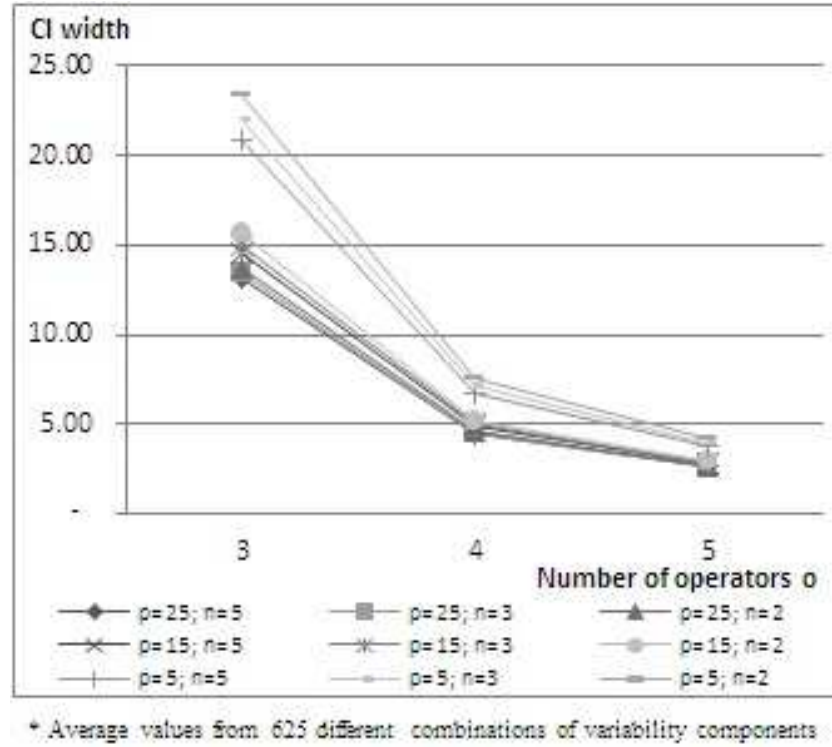


Figure 4.17: Average CI width of σ_{GRR}^2 for different number of operator o

Figure 4.18 shows that when the number of replications is the same, wider CI is obtained in a design with a small sample size and number of operators. Increasing only the sample size will cause a little reduction on the average width. But by increasing the number of operators, the width of the CI is substantially reduced. Only in the case when sample size and number operators are small, an increase on the number of replications reduces the simulated average CI width. For all the other scenarios number of replications seems to have a small effect on the reduction of the confidence interval width. As an example, combination ($p=25, o=5, n=2$)

with an average width of 2.69, presents a reduction of just 1.6% when 1 replication is increased in the design.

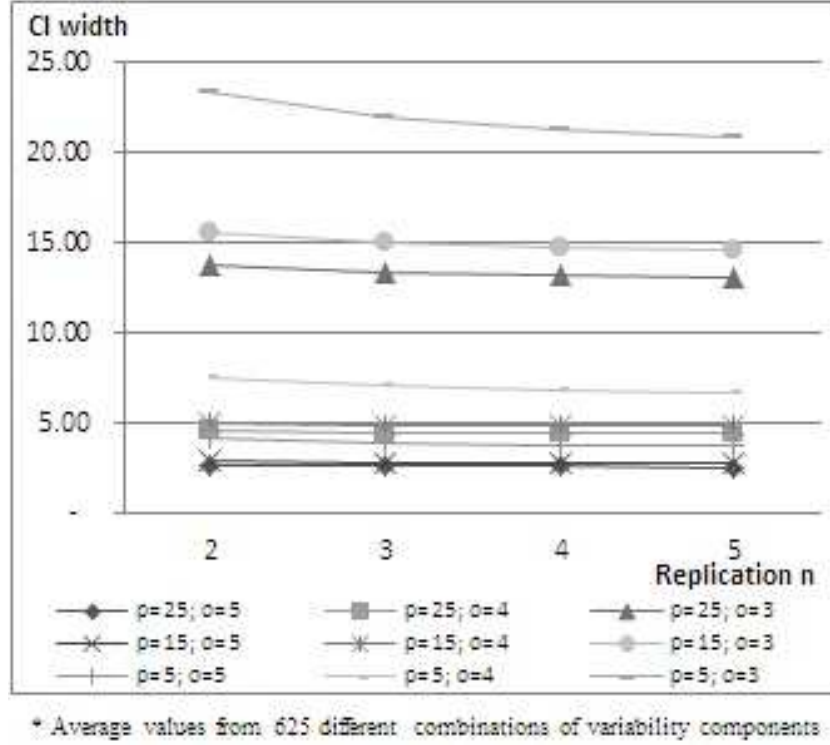


Figure 4.18: Average CI width of σ_{GRR}^2 for different replication n

Table 4.7 shows all the 25 total measurement scenarios analyzed in the present research. Parameter allocation and average CI width for the worst design and best design are shown. Last column of the table indicates the percentage of improvement passing from the worst to the best design. It can be seen that when possible, number of replications in the worst scenario is reduced at most and increased number of operators and sample size for obtaining the best scenario. For example, in the first row, for a total measurement number of 30 there is one possible design ($p=5, o=3, n=2$), so there is not improvement. A design of 150 total measurements presents as a worst design ($p=10, o=3, n=5$), which has an average width of 16.36. The best design for 150 total measurements is ($p=15, o=5, n=2$) that gives an average width of 2.92. This corresponds to a reduction of 82.14% in the average width of the

confidence interval.

Table 4.7: % reduction \overline{CI} width σ_{GRR}^2 for different designs with same measurement number

Total measurement	Worst Design				Best Design				Percentage of reduction
	p	o	n	\overline{CI} width σ_{GRR}^2	p	o	n	\overline{CI} width σ_{GRR}^2	
30	5	3	2	23.43	5	3	2	23.43	0.00%
40	5	4	2	7.49	5	4	2	7.49	0.00%
45	5	3	3	22.04	5	3	3	22.04	0.00%
50	5	5	2	4.15	5	5	2	4.15	0.00%
60	5	3	4	21.33	5	4	3	7.07	66.85%
75	5	3	5	20.91	5	5	3	3.93	81.21%
80	5	4	4	6.86	10	4	2	5.74	16.31%
90	10	3	3	17.01	15	3	2	15.62	8.18%
100	5	4	5	6.73	10	5	2	3.23	52.03%
120	10	3	4	16.60	15	4	2	5.12	69.17%
125	5	5	5	3.75	5	5	5	3.75	0.00%
135	15	3	3	15.04	15	3	3	15.04	0.00%
150	10	3	5	16.36	15	5	2	2.92	82.14%
160	10	4	4	5.40	20	4	2	4.81	11.00%
180	15	3	4	14.76	15	4	3	4.96	66.37%
200	10	4	5	5.33	20	5	2	2.77	48.01%
225	15	3	5	14.59	15	5	3	2.85	80.49%
240	20	3	4	13.80	20	4	3	4.69	66.00%
250	10	5	5	3.03	25	5	2	2.69	11.23%
300	20	3	5	13.68	20	5	3	2.72	80.12%
320	20	4	4	4.64	20	4	4	4.64	0.00%
375	25	3	5	13.12	25	5	3	2.64	79.85%
400	20	4	5	4.61	20	5	4	2.69	41.57%
500	25	4	5	4.47	25	5	4	2.62	41.28%
625	25	5	5	2.61	25	5	5	2.61	0.00%

4.1.7 Analysis of total variability σ_{Total}^2

Total variability of the system is the summation of measurement system variability and process variability. For same parameter allocation it can be noticed that the average CI width of σ_{Total}^2 is smaller than the CI average width of σ_{GRR}^2 . Similar to the average CI width of σ_{GRR}^2 , the average CI width of σ_{Total}^2 is especially influenced by the number of operators in the design. Following graphics present the individual effect of each parameter on the CI average width.

Figure 4.19 shows that when having a reduced number of operators in the design, no matter

the increment on the sample size, the average CI width remains large. But it also can be noticed that when having the lowest number of operators in the design, increasing the sample size considerable reduces the CI width. This is evident considering design (p=10,o=3,n=5) which has an average width of 11.98. Increasing the sample size to 15, the new design (p=15,o=3,n=5) presents an average width of 10.95. Adding more samples reduced the average width 8.6%.

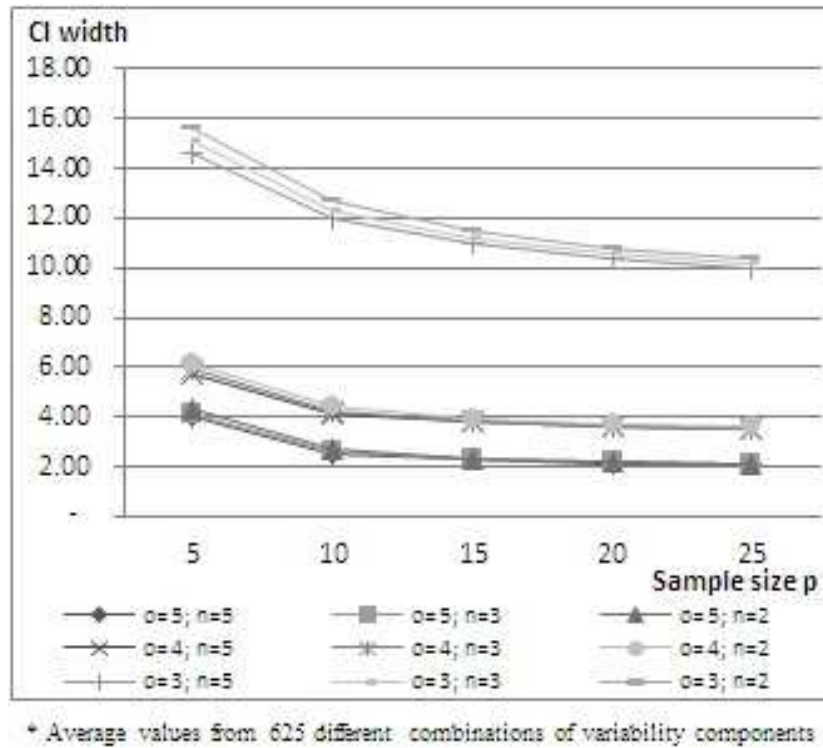


Figure 4.19: Average CI width of σ_{Total}^2 for different sample size p

On Figure 4.20 it is clearly evidenced that for any allocation of sample size and number of replications, an increment on the number of operators importantly reduces the size of the average CI width. To illustrate, combination (p=5,o=3,n=5) returns a CI average width of 14.61. And by increasing the number of operators, the average CI width of the new design (p=5,o=4,n=5) changes to 5.77. This implies a reduction of 60.5% with respect to the original average width.

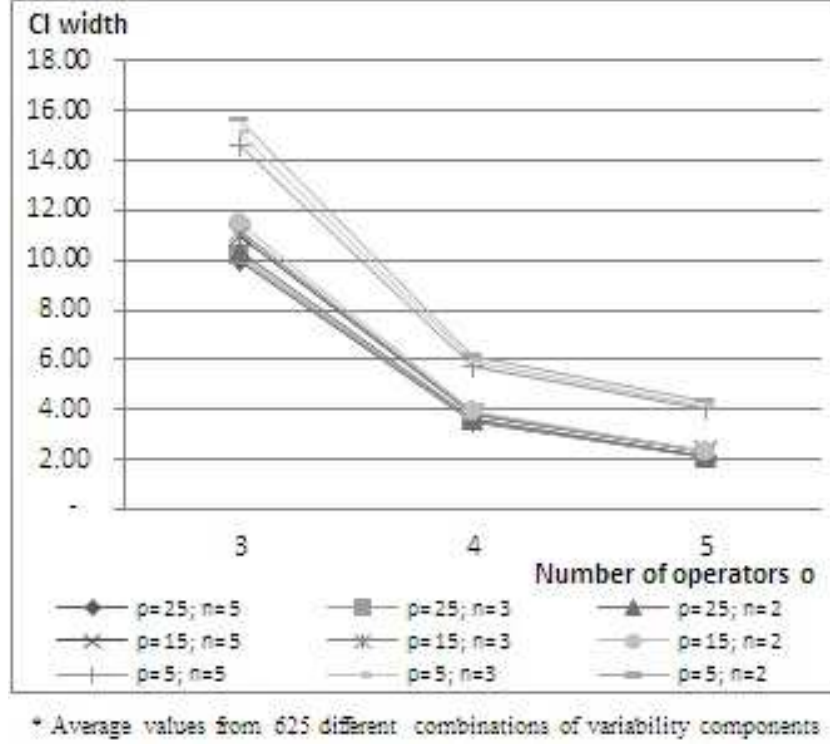


Figure 4.20: Average CI width of σ^2_{Total} for different number of operator o

Figure 4.21 shows the impact of number of replications on the average CI width. Increasing only the number of replications is not significant for reducing the average CI width. This can be noticed with the design (p=25,o=4,n=4) that has an average CI width of 3.49. Increasing 1 replication to the design changes the average width to 3.47, that is a reduction of only 0.5%.

Table 4.8 presents a comparison of the worst design and best design for different fixed number of total measurements. The designs with the worst and the best average widths are the same obtained in previous section for σ^2_{Total} . Depending on the parameter to change, reduction can or cannot be significant. In the case of having 90 total measurements, passing from the worst parameter allocation (p=10,o=3,n=3) that gives an average CI width of 12.3, to the best design (p=15,o=3,n=2) that gives an average width of 11.48, it is obtained a reduction of just 6.67%. But considering a scenario with 300 total measurements, passing

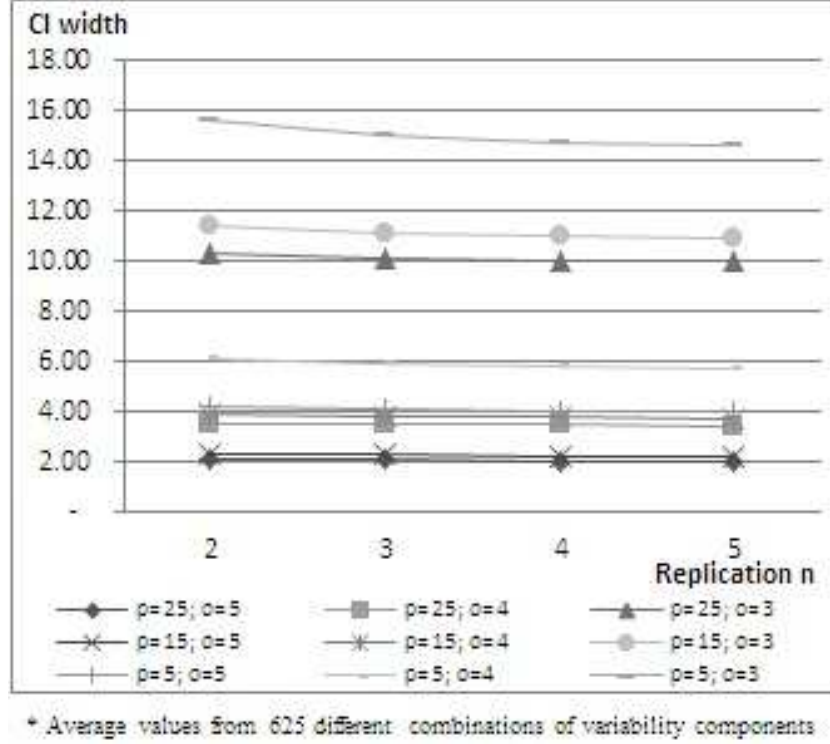


Figure 4.21: Average CI width of σ_{Total}^2 for different replication n

from the worst design ($p=20, o=3, n=5$) that gives an average width of 10.37 to the best design ($p=20, o=5, n=3$) that gives an average width of 2.17, then the reduction represents 79.06% of the original value.

4.2 Analysis of measurement system capability metrics

This section presents results of the simulation analysis performed on the most common measurement system capability metrics. All of these metrics are functions of the variability components analyzed in last section, so it is expected a similar behavior on the findings. It is given an objective criterion about the best Gauge R&R study designs for each capability metric. Parameter allocation should be selected depending on the metric used on industry.

Table 4.8: % reduction \overline{CI} width σ_{Total}^2 for different designs with same measurement number

Total measurement	Worst Design				Best Design				Percentage of reduction
	p	o	n	\overline{CI} width σ_{Total}^2	p	o	n	\overline{CI} width σ_{Total}^2	
30	5	3	2	15.66	5	3	2	15.66	0.00%
40	5	4	2	6.18	5	4	2	6.18	0.00%
45	5	3	3	15.08	5	3	3	15.08	0.00%
50	5	5	2	4.28	5	5	2	4.28	0.00%
60	5	3	4	14.79	5	4	3	5.96	59.70%
75	5	3	5	14.61	5	5	3	4.13	71.76%
80	5	4	4	5.84	10	4	2	4.39	24.80%
90	10	3	3	12.30	15	3	2	11.48	6.67%
100	5	4	5	5.77	10	5	2	2.67	53.82%
120	10	3	4	12.10	15	4	2	3.94	67.47%
125	5	5	5	4.00	5	5	5	4.00	0.00%
135	15	3	3	11.19	15	3	3	11.19	0.00%
150	10	3	5	11.98	15	5	2	2.35	80.38%
160	10	4	4	4.21	20	4	2	3.71	11.85%
180	15	3	4	11.04	15	4	3	3.85	65.15%
200	10	4	5	4.17	20	5	2	2.21	47.02%
225	15	3	5	10.95	15	5	3	2.30	78.96%
240	20	3	4	10.45	20	4	3	3.64	65.17%
250	10	5	5	2.54	25	5	2	2.13	16.32%
300	20	3	5	10.37	20	5	3	2.17	79.06%
320	20	4	4	3.61	20	4	4	3.61	0.00%
375	25	3	5	10.00	25	5	3	2.10	79.04%
400	20	4	5	3.58	20	5	4	2.15	39.90%
500	25	4	5	3.47	25	5	4	2.08	40.00%
625	25	5	5	2.07	25	5	5	2.07	0.00%

4.2.1 Analysis of part variability to gauge variability ratio ρ_p

Graphics below show the effect of different parameter allocations on the average confidence interval width of ρ_p . It can be observed in Figure 4.22 that independently the number of operators or number of replications in the design of the Gauge R&R study, increasing the number of samples in the experiment results in an important reduction of the CI width, especially when having a small sample size. For example, parameter allocation (p=5,o=3,n=5) has an average CI width of 10.73. By increasing 5 samples, the new design (p=10,o=3,n=5) presents an average width of 4.97, that means a reduction of 53.7%. Now, if considering a design with a larger number of samples, like (p=20,o=3,n=5), the width of the CI obtained

is 3.28. And by adding 5 samples to previous design, the average width decreased to 2.96, which is a reduction of only 9.6%.

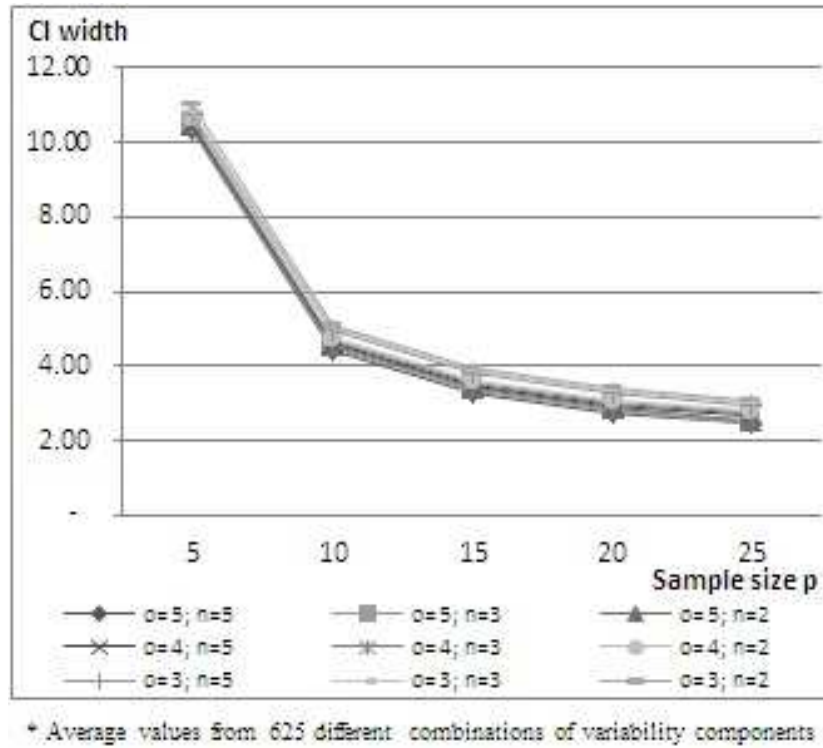


Figure 4.22: Average CI width of ρ_p for different sample size p

Figure 4.23 points out that the three scenarios with a reduced number of parts in the design, ($p=5, n=5$), ($p=5, n=3$) and ($p=5, n=2$), give a CI more than three times wider than the other six scenarios. For example, considering the maximum number of operators and number of replications, a Gauge R&R study with parameter allocation ($p=5, o=5, n=5$) has a average CI width of 10.32, that is 3.2 times wider than the average width of the design ($p=15, o=5, n=5$). This indicates that no matter the number of operators neither number of replications, having a small sample size seriously increase the width of the confidence interval. For all the scenarios it is noticed that increasing the number of operators have a moderate effect on the reduction of the average confidence interval width. As an example, with a design ($p=15, o=3, n=2$) the average CI width obtained is 3.93. Assigning an extra operator to the design changes the

average width to 3.63. This represents a reduction of 7.6%.

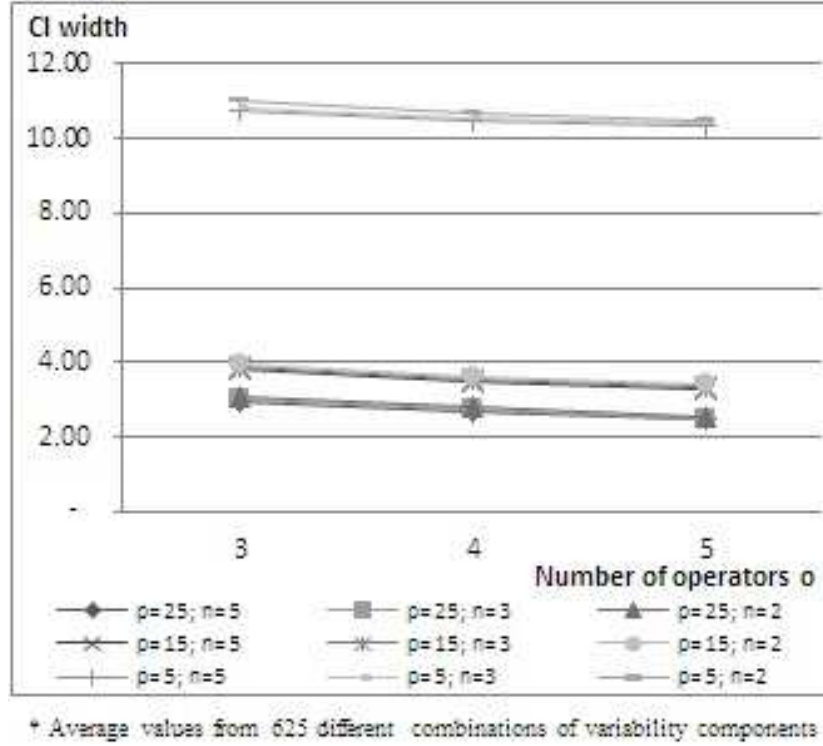


Figure 4.23: Average CI width of ρ_p for different number of operator o

Similar to previous graphic, Figure 4.24 shows how designs with low level of sample size have a wider CI than other scenarios. It is also noticed that increasing number of replications has practically no effect on the average width of the confidence interval. For example, considering a design with parameter allocation ($p=25, o=4, n=2$) produces an average CI width of 2.78. By increasing 1 replication, the new design ($p=25, o=4, n=3$) returns an average width of 2.72. This is a decrement of only 2.2%.

Table 4.9 presents all possible parameter allocations for three fixed total measurements in a Gauge R&R study. It can be observed that designs with a large number of parts tend to hold narrower confidence intervals. Looking at the designs with 300 measurements and variability components $\sigma_{Rep}^2 = 1$, $\sigma_p^2 = 3$, $\sigma_o^2 = 3$ and $\sigma_{op}^2 = 5$, design ($p=25, o=4, n=3$) has an average CI width of 2.83, the smallest from all possible parameter allocation combinations.

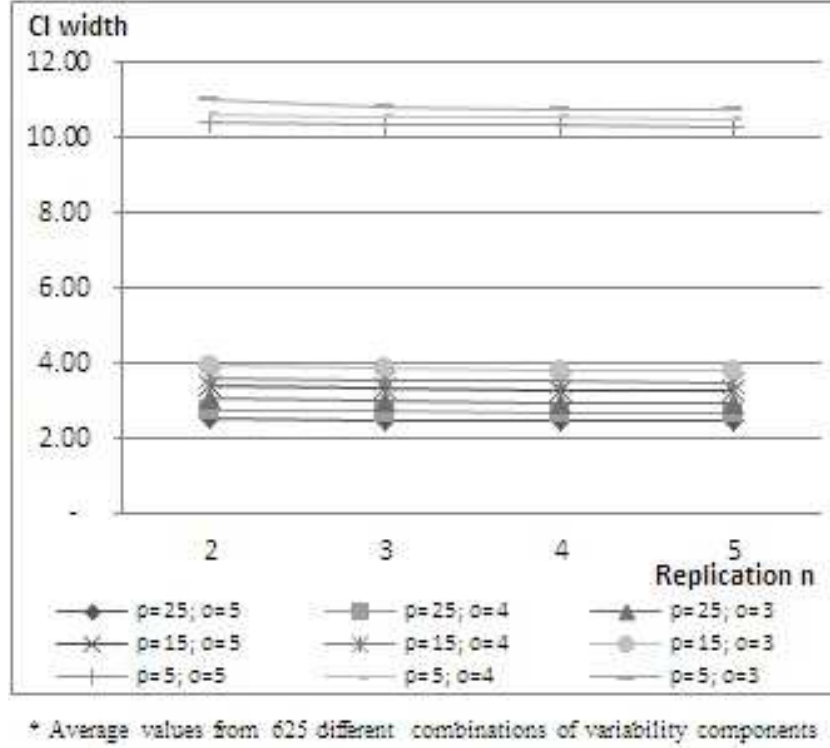


Figure 4.24: Average CI width of ρ_p for different replication n

It is important to remark that although increasing sample size in the design is a priority for reducing the average width, a consequent serious reduction on the number of operators may have an adverse effect. Considering the same variability components, design $(p=20, o=5, n=3)$ with an average width of 2.91 presents an improvement compared to design $(p=25, o=3, n=4)$ with an average width of 3.18.

4.2.2 Analysis of proportion of measurement system variability due to repeatability ρ_{Rep}

Average confidence interval width of proportion of measurement system variability due to repeatability ρ_{Rep} exhibits a similar behavior when modifying any of the three parameters. From Figure 4.25 it can be seen that, independently of sample size, designs with largest number of operators, i.e. $(o=5, n=5), (o=5, n=3)$ and $(o=5, n=2)$, have the smallest average CI width. It is noticed that regardless the number of operators and number of replications

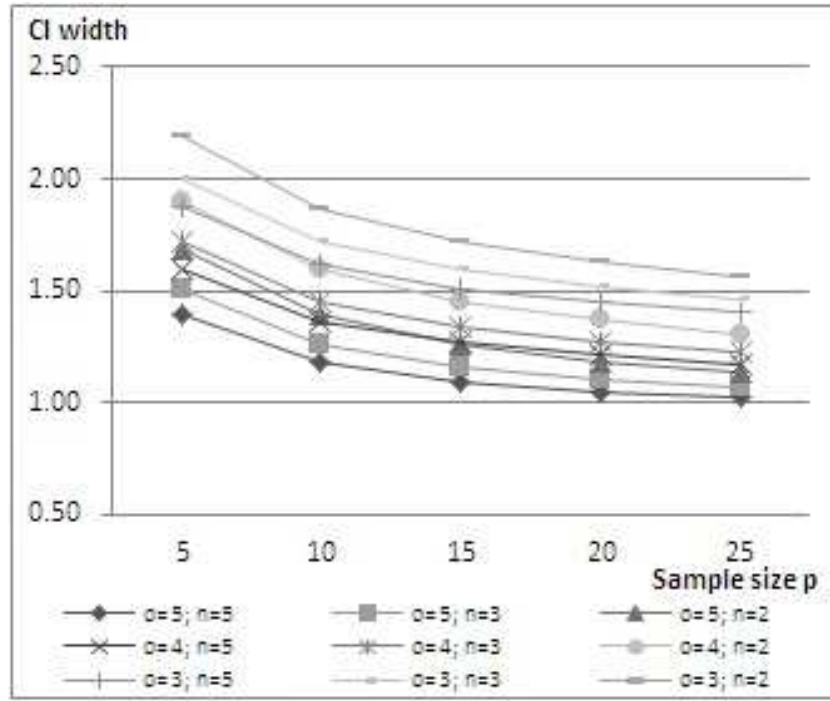
Table 4.9: Average CI width of ρ_p for different total measurement numbers

				σ_{Rep}^2	1	1	1	3	3	3	5	5	5
				σ_p^2	1	5	3	1	5	3	1	5	3
				σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5	
120	10	3	4	5.02	4.62	5.29	5.12	4.98	5.32	4.59	4.70	5.21	
	10	4	3	4.73	4.24	4.96	4.93	4.60	5.04	4.53	4.41	4.98	
	15	4	2	3.62	3.08	3.77	3.90	3.50	3.91	3.57	3.33	3.89	
	20	3	2	3.41	2.87	3.55	3.65	3.34	3.69	3.24	3.12	3.63	
200	10	4	5	4.65	4.21	4.94	4.87	4.53	4.99	4.45	4.31	4.91	
	10	5	4	4.41	3.95	4.67	4.69	4.26	4.76	4.36	4.09	4.71	
	20	5	2	2.80	2.33	2.93	3.19	2.72	3.09	2.92	2.55	3.11	
	25	4	2	2.72	2.24	2.85	3.11	2.66	3.02	2.76	2.47	3.02	
300	15	4	5	3.46	3.02	3.73	3.73	3.37	3.80	3.37	3.14	3.74	
	15	5	4	3.22	2.80	3.46	3.53	3.13	3.56	3.26	2.94	3.53	
	20	3	5	3.26	2.82	3.51	3.52	3.20	3.59	3.08	2.94	3.50	
	20	5	3	2.72	2.31	2.91	3.08	2.67	3.04	2.81	2.48	3.03	
	25	3	4	2.95	2.50	3.18	3.25	2.90	3.28	2.83	2.65	3.21	
	25	4	3	2.65	2.22	2.83	3.01	2.60	2.97	2.68	2.39	2.94	

in the design, increasing sample size has a moderate effect on the reduction of the average CI width. Regarding the case with parameter allocation ($p=10, o=5, n=2$), the resulting average width is 1.39. By increasing 5 samples, the average width of the new design ($p=15, o=5, n=2$) changes to 1.26. It signifies a decrease of 9.4%.

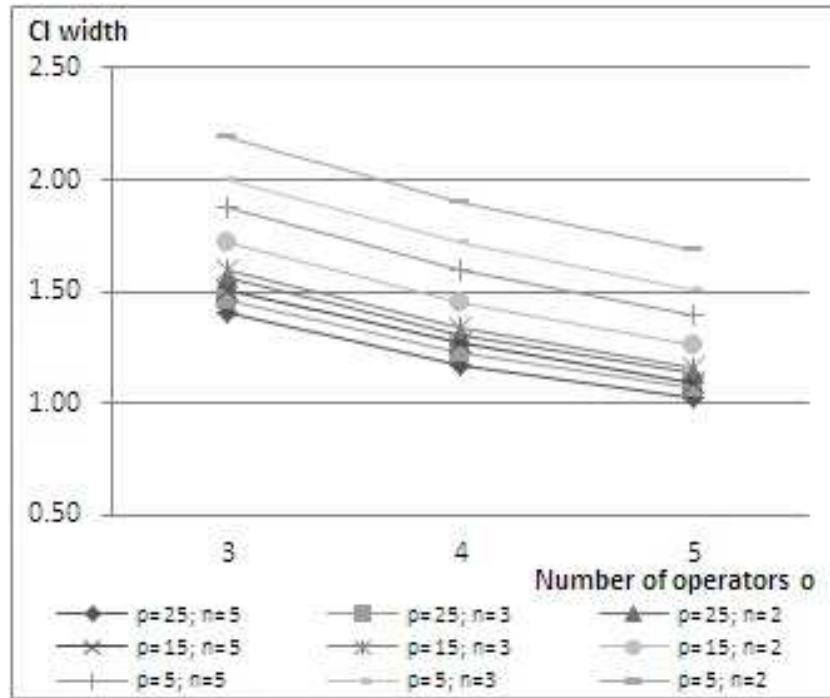
For Figure 4.26, it is evident that when having the same number of operators in the design, increasing sample size instead of number of replications results on a larger reduction of the average CI width. As an example, parameter allocation ($p=15, o=4, n=3$) with average width 1.34 results in a better option than design ($p=5, o=4, n=5$) with average width 1.60. Modifying the number of operators in the study has a moderate effect on the width of the confidence interval. To illustrate this, design ($p=5, o=4, n=2$) presents an average width of 1.90, while design ($p=5, o=5, n=2$) returns an average width of 1.68. This implies a decrease of 11.3%.

Figure 4.27 shows that when having the same number of replications, increasing any of both,



* Average values from 625 different combinations of variability components

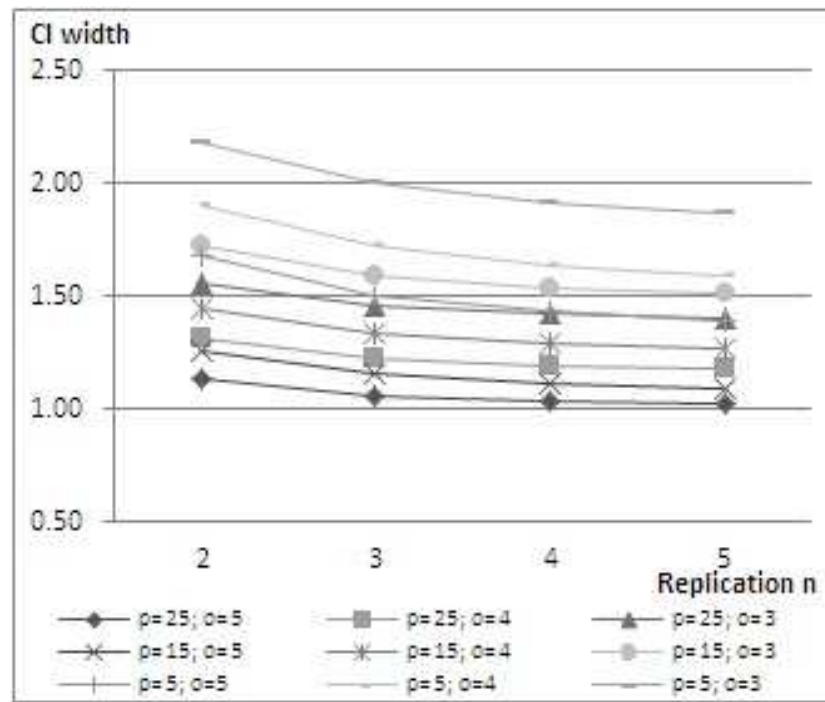
Figure 4.25: Average CI width of ρ_{Rep} for different sample size p



* Average values from 625 different combinations of variability components

Figure 4.26: Average CI width of ρ_{Rep} for different number of operator o

number of operators or number of replications, is significant for reducing the average CI width. For parameter allocation combination ($p=5, o=4, n=3$), where sample size is small, the average width is 1.72. By increasing the number of parts and reducing number of operators, a new design ($p=15, o=3, n=3$) results on an average CI width of 1.59. Similar to previous graphics, it is noticed that no matter the number of operators and sample size in the design of the Gauge R&R study, increasing the number of replications has a moderate effect on the reduction of the average CI width. As an example, passing from a design with a reduced number of replications ($p=25, o=5, n=2$) to one with larger number of replications ($p=25, o=5, n=3$), changes the average width from 1.14 to 1.07. Meaning a reduction of 6.5%.



* Average values from 625 different combinations of variability components

Figure 4.27: Average CI width of ρ_{Rep} for different replication n

It is important to select the proper parameter allocation combination when having a limited total number of measurements in order to guarantee a narrower confidence interval. From Table 4.10, it is shown that designs with largest number of operators result on the smallest

average CI widths. Also it is noticed that when having the maximum number of operators in the design, it is preferable to increase sample size instead of number of replications. Following example illustrates previous observations. Having the scenario with $\sigma_{Rep}^2 = 1$, $\sigma_p^2 = 5$, $\sigma_o^2 = 1$, $\sigma_{op}^2 = 3$ and 200 total measurements, design (p=10,o=5,n=4) presents an average width of 1.19, and design (p=20,o=5,n=2) an average width of 1.11. This are the smallest values from the four possible different parameter allocation combinations.

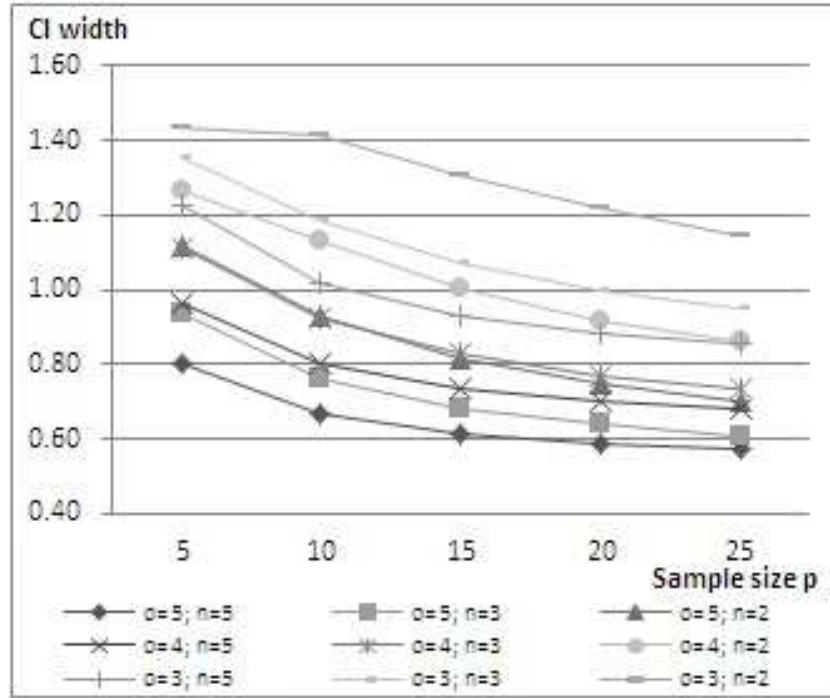
Table 4.10: Average CI width of ρ_{Rep} for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	1.65	1.72	1.91	1.57	1.79	1.79	1.16	1.50	1.64
	10	4	3	1.45	1.48	1.65	1.40	1.58	1.57	0.98	1.31	1.45
	15	4	2	1.46	1.44	1.61	1.40	1.60	1.57	0.96	1.32	1.46
	20	3	2	1.63	1.62	1.79	1.55	1.78	1.75	1.13	1.49	1.63
200	10	4	5	1.35	1.38	1.55	1.34	1.49	1.48	0.92	1.22	1.35
	10	5	4	1.20	1.19	1.36	1.21	1.34	1.31	0.78	1.07	1.20
	20	5	2	1.19	1.11	1.28	1.23	1.34	1.28	0.76	1.07	1.19
	25	4	2	1.31	1.23	1.40	1.34	1.46	1.40	0.89	1.20	1.31
300	15	4	5	1.25	1.23	1.40	1.29	1.40	1.37	0.86	1.13	1.25
	15	5	4	1.11	1.05	1.23	1.17	1.26	1.21	0.72	0.99	1.11
	20	3	5	1.43	1.42	1.59	1.45	1.58	1.54	1.06	1.31	1.43
	20	5	3	1.09	1.01	1.19	1.17	1.26	1.19	0.72	0.98	1.10
	25	3	4	1.41	1.37	1.54	1.44	1.56	1.51	1.05	1.30	1.41
	25	4	3	1.22	1.14	1.31	1.29	1.37	1.31	0.85	1.11	1.22

4.2.3 Analysis of proportion of measurement system variability due to reproducibility $\rho_{Reproducibility}$

The average confidence interval width of $\rho_{Reproducibility}$ tends to be affected by changes of any of the three parameters under study. Figure 4.28 reveals that augmenting sample size produces a moderate reduction on the width of the CI for all combinations of number of operators and number of replications observed. Previous remark can be evidenced considering parameter allocation combination (p=5,o=4,n=5) which has an average width of 0.96. Rising sample

size to 10 in the design modifies the average width to 0.80. This represents a reduction of 16.8%.



* Average values from 625 different combinations of variability components

Figure 4.28: Average CI width of $\rho_{Reproducibility}$ for different sample size p

From Figure 4.29 it can be also concluded that regardless other parameter combinations, an increase on number of operators has a moderate effect on the reduction of the average CI width. For example, moving from a design ($p=15, o=4, n=5$) with an average width of 0.74, to a design ($p=15, o=5, n=5$) with an average width of 0.62, occasions a decrease of 16.2%.

Number of replications moderately affects the average CI width no matter the allocation of number of operators and sample size. As an example from Figure 4.30, it can be considered the parameter allocation combination ($p=5, o=3, n=3$) with an average width of 1.36. Increasing the number of replications to 4, reduces the average width of the new design ($p=5, o=3, n=4$) to 1.28, that is a decrease of 5.8%.

From Table 4.11, it can be seen that when having a fixed number of measurements, it is

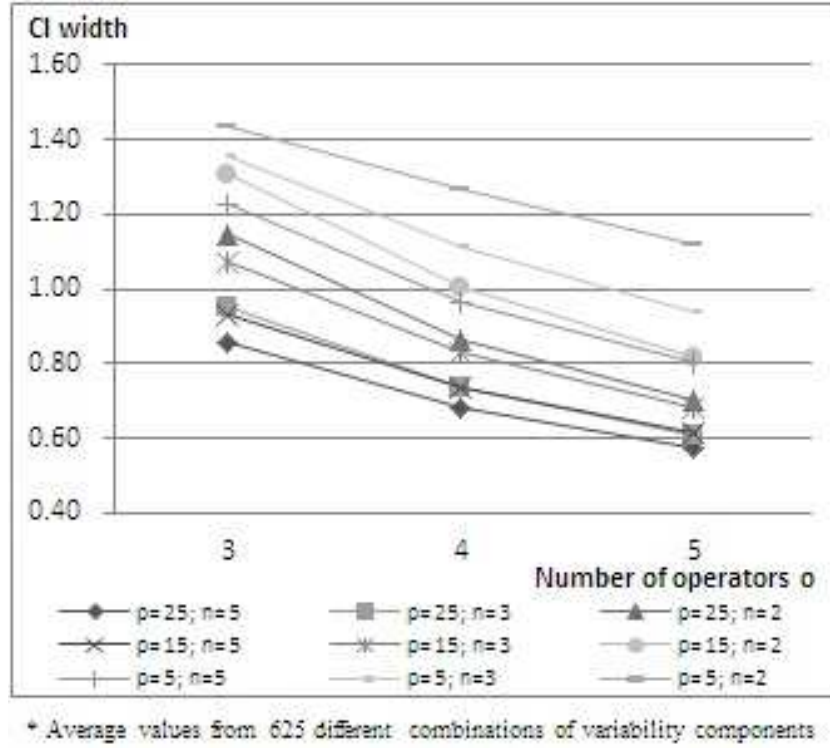


Figure 4.29: Average CI width of $\rho_{Reproducibility}$ for different number of operator o

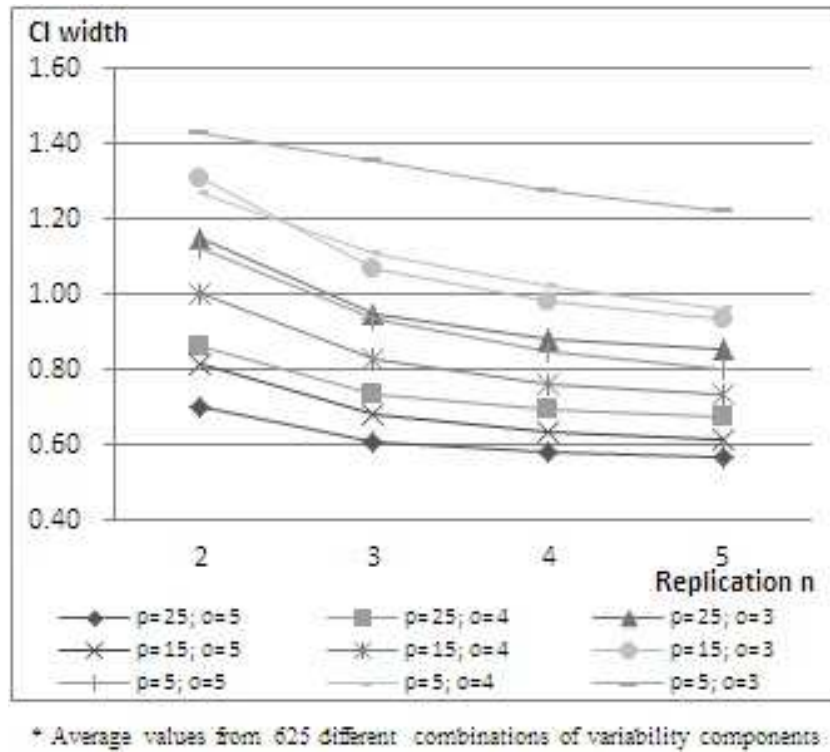


Figure 4.30: Average CI width of $\rho_{Reproducibility}$ for different replication n

recommended to increase as much as possible the number of operators in the design. Taking as an example the scenario with variability components $\sigma_{Rep}^2 = 5$, $\sigma_p^2 = 3$, $\sigma_o^2 = 5$, $\sigma_{op}^2 = 5$ and 120 total measurements, designs with 4 operators (p=10,o=4,n=3) and (p=15,o=4,n=2) produce narrower confidence intervals, 0.79 and 0.83 respectively.

Table 4.11: Average CI width of $\rho_{Reproducibility}$ for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	0.89	0.41	0.25	1.80	0.75	0.60	4.33	1.38	0.94
	10	4	3	0.77	0.36	0.21	1.51	0.66	0.51	3.60	1.21	0.79
	15	4	2	0.80	0.35	0.21	1.70	0.69	0.51	3.98	1.36	0.83
	20	3	2	0.93	0.39	0.23	2.17	0.79	0.58	5.04	1.64	0.98
200	10	4	5	0.70	0.33	0.20	1.26	0.61	0.47	2.94	1.04	0.72
	10	5	4	0.62	0.29	0.17	1.10	0.55	0.42	2.47	0.92	0.63
	20	5	2	0.63	0.27	0.16	1.24	0.56	0.41	3.11	0.98	0.63
	25	4	2	0.70	0.30	0.18	1.47	0.62	0.45	3.87	1.11	0.71
300	15	4	5	0.66	0.30	0.18	1.17	0.58	0.44	2.50	0.98	0.67
	15	5	4	0.58	0.26	0.16	1.02	0.52	0.38	2.09	0.85	0.58
	20	3	5	0.76	0.34	0.20	1.42	0.67	0.50	3.41	1.15	0.78
	20	5	3	0.57	0.25	0.15	1.03	0.51	0.38	2.26	0.85	0.57
	25	3	4	0.75	0.33	0.20	1.43	0.66	0.49	3.62	1.14	0.77
	25	4	3	0.64	0.28	0.17	1.19	0.57	0.42	2.85	0.97	0.65

4.2.4 Analysis of precision-to-tolerance ratio P/T

Confidence interval of precision-to-tolerance ratio P/T is moderately affected by a change in the sample size as shown on Figure 4.31. Going from a parameter allocation combination (p=15,o=4,n=2) with an average width of 1.54, to a design (p=20,o=4,n=2) with an average width of 1.46, results on a reduction of 5.1%. It is also noticed that for same sample size, impact on the wide of the average width is due only to number of operators. It is clearly shown in the graphic that scenarios with largest number of operators (o=5,n=5), (o=5,n=3) and (o=5,n=2) results on the smallest average widths regardless the number of parts in the

design.

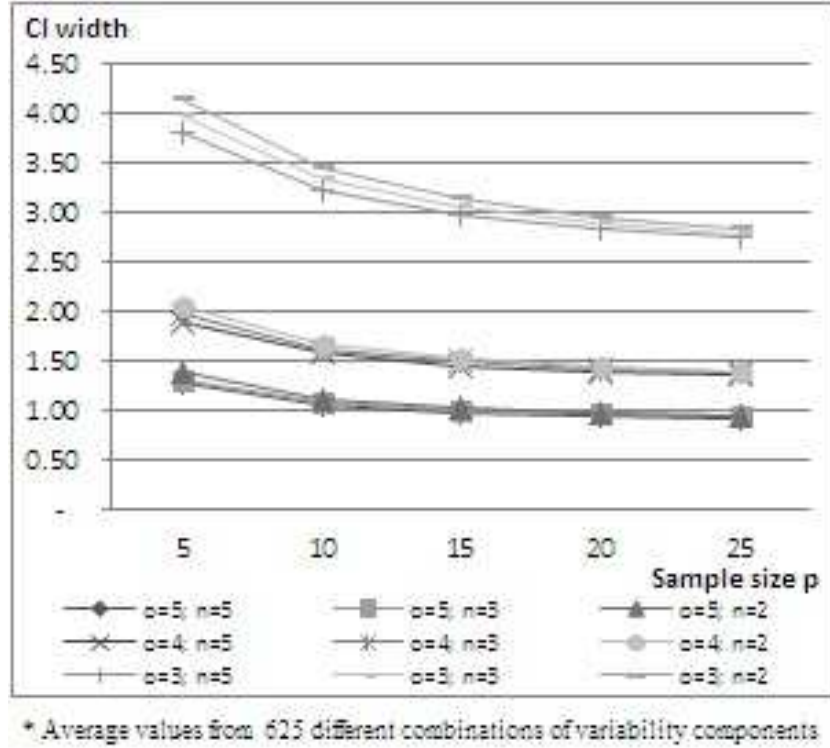


Figure 4.31: Average CI width of P/T for different sample size p

It is evidenced in Figure 4.32 that an increment on the number of operators has a high effect on the reduction of the average CI width, especially when considering designs with reduced number of operators, i.e. 3 operators. As an illustration, design $(p=5, o=4, n=5)$ produces an average width of 1.90. By increasing number of operators to 5, the new design $(p=5, o=5, n=5)$ changes its average width to 1.27, a decrease of 32.9% with respect to the original value. From the graphic, it can be also pointed out that when having the same number of operators, a design with more samples is better than one with more replications. With respect to number of replications, a change on this parameter has virtually no effect on the average CI width. Considering Figure 4.33, for the case where the design is $(p=15, o=3, n=4)$, the resulting average confidence interval width is 3.01. And by increasing to 5 the number of replications, the average width of the design $(p=15, o=3, n=5)$ changes

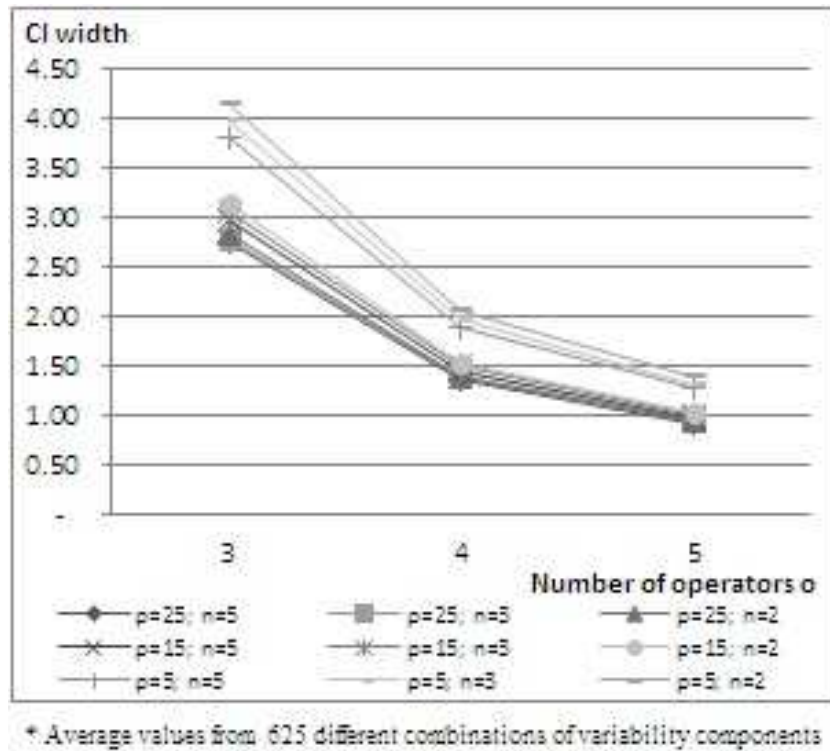


Figure 4.32: Average CI width of P/T for different number of operator o

only to 2.98 This is an insignificant reduction of just 0.9%.

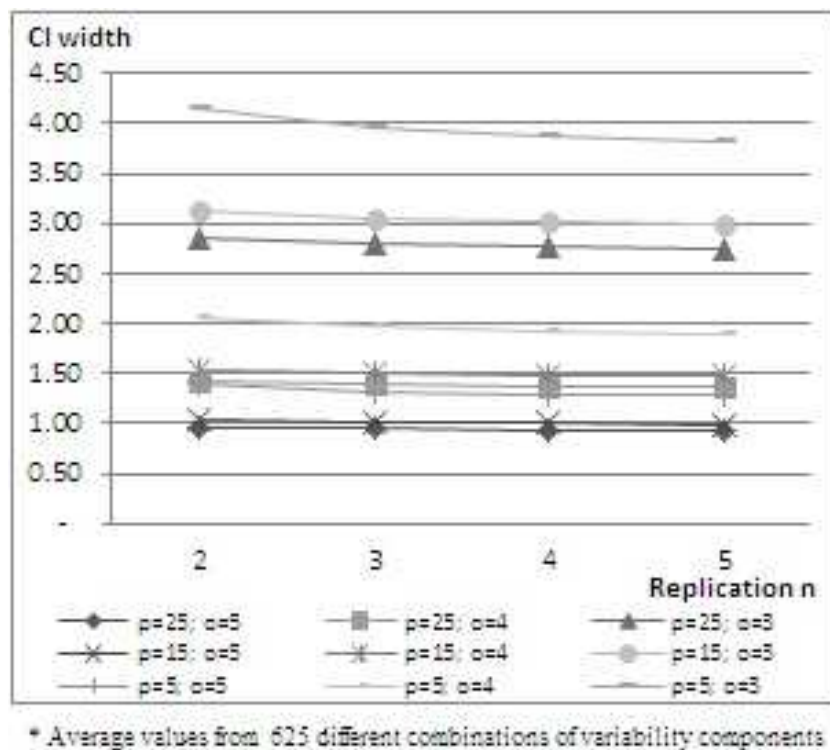


Figure 4.33: Average CI width of P/T for different replication n

Table 4.12 helps to identify the best design when having a limited number of measurements in a Gauge R&R study. It can be observed that designs with larger number of operators tend to reduce the average width to its minimum values, and in second place sample size can also be considered for its reduction. Number of replications has not a notorious effect and can be minimized in order to increase any of the previous parameters. For example, scenario with variability components $\sigma_{Rep}^2 = 3$, $\sigma_p^2 = 5$, $\sigma_o^2 = 5$, $\sigma_{op}^2 = 3$, and 200 total measurements reaches the smallest average width 1.19 with the design (p=20,o=5,n=2). This design reduces to its minimum the number of replications to increase to its highest value the number of operators and also considers a large sample size.

Table 4.12: Average CI width of P/T for different total measurement numbers

Measurements	p	o	σ_{op}^2 n	σ_{Rep}^2	1	1	1	3	3	3	5	5	5
				σ_p^2	1	5	3	1	5	3	1	5	3
				σ_o^2	1	1	3	3	5	5	1	3	5
					1	3	5	1	3	5	1	3	5
120	10	3	4		3.28	3.17	3.59	3.35	3.58	3.53	2.29	2.99	3.29
	10	4	3		1.63	1.52	1.76	1.72	1.83	1.77	1.11	1.46	1.63
	15	4	2		1.53	1.34	1.60	1.69	1.76	1.66	1.07	1.39	1.54
	20	3	2		2.97	2.62	3.09	3.23	3.34	3.19	2.16	2.71	2.99
200	10	4	5		1.58	1.49	1.75	1.67	1.80	1.73	1.00	1.41	1.59
	10	5	4		1.07	0.98	1.17	1.16	1.24	1.18	0.67	0.94	1.07
	20	5	2		0.99	0.80	1.02	1.13	1.19	1.08	0.63	0.87	0.99
	25	4	2		1.41	1.16	1.46	1.60	1.68	1.54	0.92	1.26	1.42
300	15	4	5		1.46	1.30	1.58	1.6	1.71	1.61	0.89	1.29	1.47
	15	5	4		1.00	0.85	1.06	1.12	1.19	1.10	0.59	0.87	1.00
	20	3	5		2.86	2.56	3.05	3.05	3.25	3.10	1.84	2.55	2.85
	20	5	3		0.97	0.79	1.01	1.11	1.17	1.07	0.58	0.85	0.97
	25	3	4		2.78	2.42	2.94	3.01	3.20	3.02	1.80	2.49	2.79
	25	4	3		1.39	1.14	1.45	1.57	1.66	1.52	0.85	1.23	1.39

4.2.5 Analysis of signal-to-noise ratio SNR

Confidence interval of signal-to-noise ratio is highly affected by sample size. Figure 4.34 shows that no matter the allocation of number of operators and number of replications, having more

parts in the Gauge R&R study significantly reduces the width of the confidence interval, especially for small sample sizes, i.e. $n=5,10$. It is shown that when having a parameter allocation combination ($p=10,o=3,n=2$) the average CI width is 2.11. By increasing the sample size to 15, the new design gets an average width of 1.83. This represents a reduction of 13.6%. When sample size is the same, a change in any of the other two parameters does not significantly affect the average width.

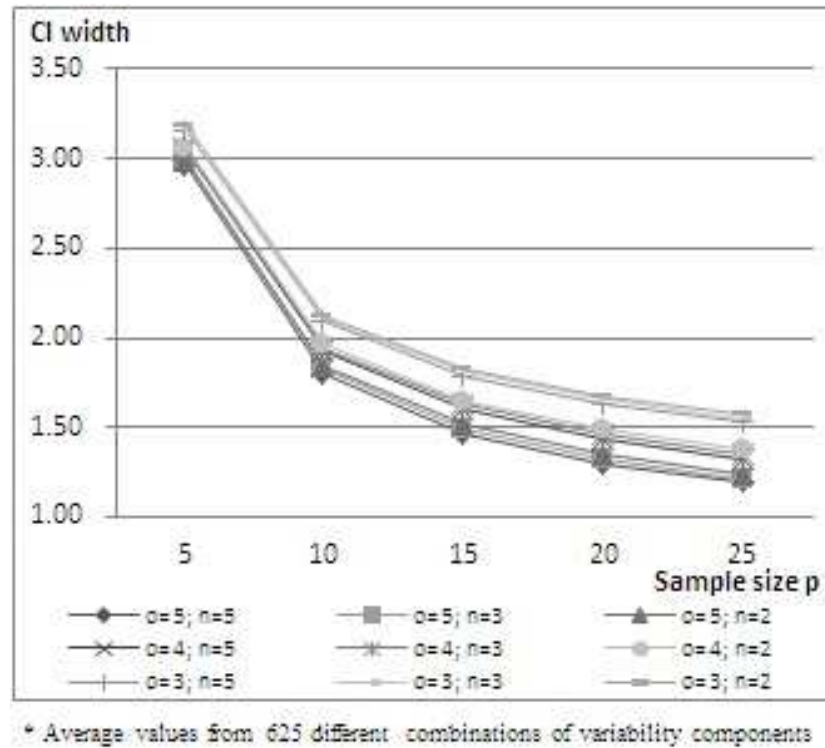


Figure 4.34: Average CI width of SNR for different sample size p

According to Figure 4.35, it is demonstrated that when having the same number of operators in the Gauge R&R study, increasing number of samples is a better option than increasing number of replications in the design. It is evident that designs with largest number of parts ($p=25,n=5$), ($p=25,n=3$) and ($p=25,n=2$) return narrower confidence intervals. The graphic also shows that increasing number of operators in the design has a moderate effect in the reduction of the average CI width. As an example, parameter allocation combination

($p=15, o=3, n=2$) gives an average width of 1.83. An increase to 4 operators, changes the average width of the design ($p=15, o=4, n=2$) to 1.65. This signifies a reduction of 9.7%.

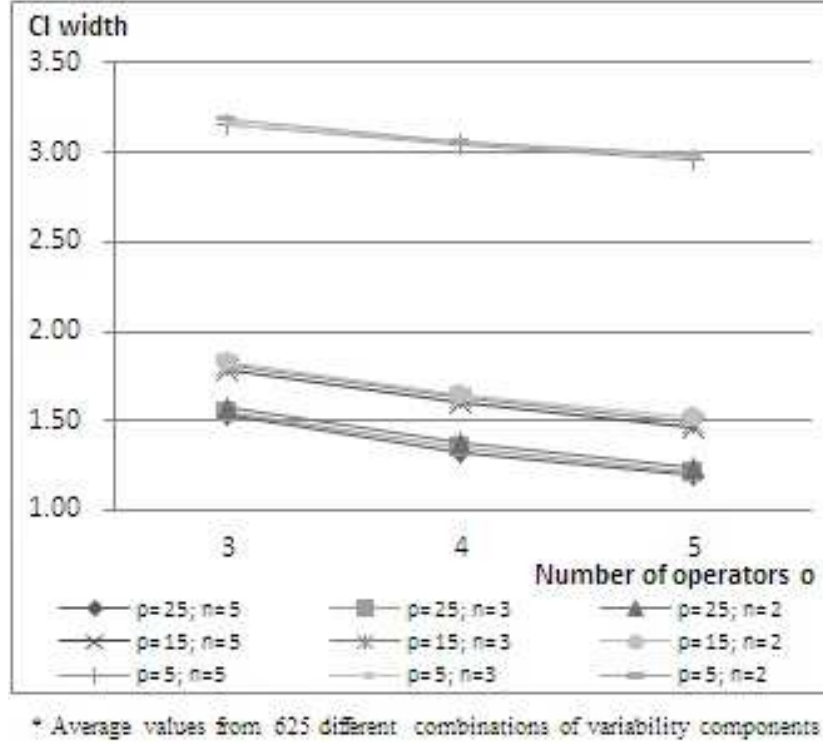


Figure 4.35: Average CI width of SNR for different number of operator o

Figure 4.36 presents the effect of number of replications on the width of the confidence interval. It is clearly evidenced that increasing number of replications in any design has a low effect on the average CI width. For example, parameter allocation combination ($p=5, o=3, n=2$) results in an average width of 3.18. Increasing 1 replication to previous design modifies it to 3.16, implying a reduction of only 0.8%.

When having a limited number of experiments to perform, it is important to select the correct level for each parameter. Table 4.13 presents all possible parameter allocation combinations for the same total measurement numbers. It is noticed that narrower confidence intervals are obtained when the design involves firstly a large sample size, and secondly a large number of operators. Number of replications are not that significant to reduce the average CI width.

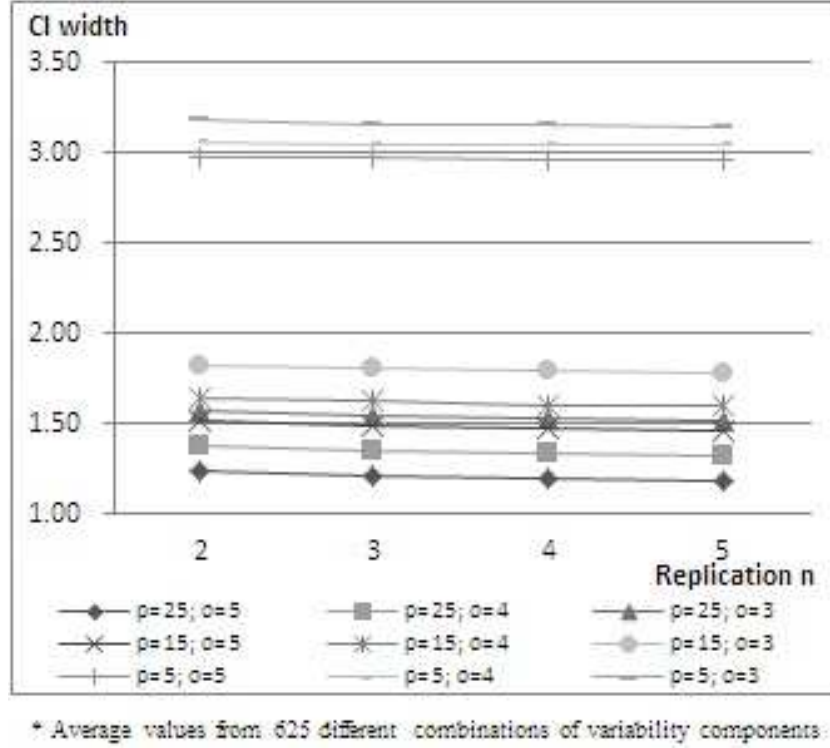


Figure 4.36: Average CI width of SNR for different replication n

To illustrate, scenario with 200 total measurements and variability components $\sigma_{Rep}^2 = 5$, $\sigma_p^2 = 5$, $\sigma_o^2 = 3$ and $\sigma_{op}^2 = 3$, produces the smallest confidence interval with the parameter allocation combination ($p=20, o=5, n=2$). This design has an average width of 1.20, that represents a reduction of 33.0% compared with the worst design.

4.3 Analysis of process capability metrics

Although main objective of the present research is to analyze the effect of Gauge R&R study design on the average confidence intervals width of variability components and measurement system capability metrics, it is also of interest to complement this analysis with some of the criteria commonly used to check capability of the manufacturing process.

Table 4.13: Average CI width of SNR for different total measurement numbers

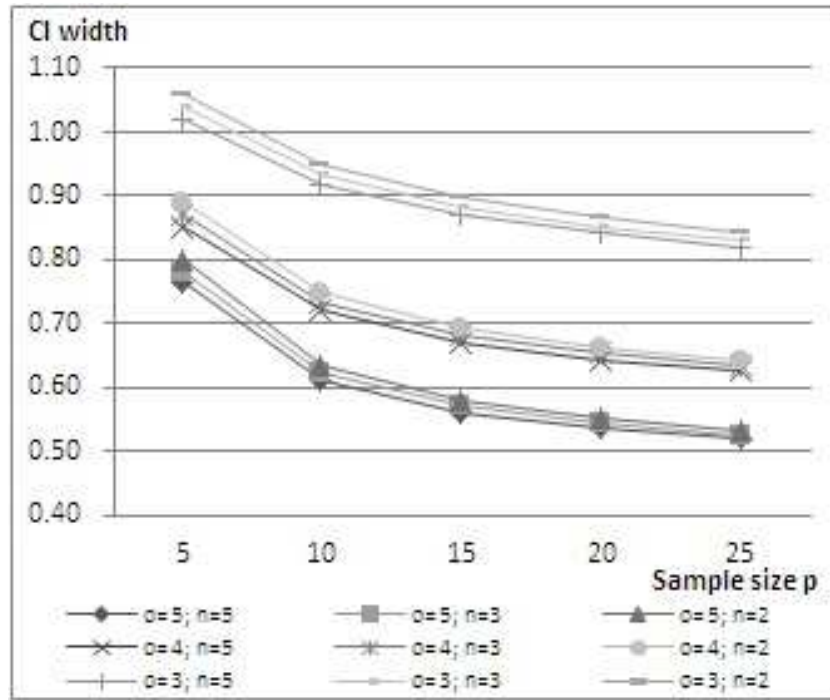
				σ_{Rep}^2	1	1	1	3	3	3	5	5	5
				σ_p^2	1	5	3	1	5	3	1	5	3
				σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5	
120	10	3	4	2.11	1.98	2.19	2.12	2.08	2.20	1.94	2.00	2.17	
	10	4	3	1.94	1.76	2.03	2.01	1.89	2.06	1.85	1.83	2.04	
	15	4	2	1.64	1.42	1.71	1.75	1.59	1.76	1.62	1.53	1.76	
	20	3	2	1.68	1.47	1.74	1.75	1.64	1.79	1.61	1.57	1.78	
200	10	4	5	1.92	1.75	2.02	1.99	1.86	2.04	1.84	1.79	2.02	
	10	5	4	1.79	1.61	1.89	1.89	1.73	1.92	1.78	1.67	1.91	
	20	5	2	1.31	1.11	1.37	1.48	1.28	1.44	1.37	1.20	1.46	
	25	4	2	1.34	1.12	1.41	1.51	1.32	1.48	1.37	1.23	1.49	
300	15	4	5	1.58	1.40	1.69	1.68	1.54	1.72	1.55	1.44	1.70	
	15	5	4	1.44	1.26	1.55	1.57	1.41	1.58	1.46	1.32	1.58	
	20	3	5	1.62	1.44	1.72	1.72	1.59	1.75	1.56	1.49	1.72	
	20	5	3	1.28	1.10	1.36	1.43	1.26	1.42	1.32	1.17	1.42	
	25	3	4	1.52	1.33	1.62	1.64	1.49	1.65	1.48	1.39	1.63	
	25	4	3	1.31	1.12	1.40	1.47	1.29	1.45	1.33	1.19	1.45	

4.3.1 Analysis of observed potential capability of the process C_p

Graphics below show the influence of the three parameters under study over the average C_p confidence interval width. Figure 4.37 clearly shows that augmenting sample size has a moderate effect on the reduction of the average CI width. Considering the case with parameter allocation combination (p=15,o=4,n=3) the confidence interval presents a width of 0.68. By increasing 5 samples, the new design (p=20,o=4,n=3) results on an average CI width of 0.65, meaning a reduction of 4.3% from the original value.

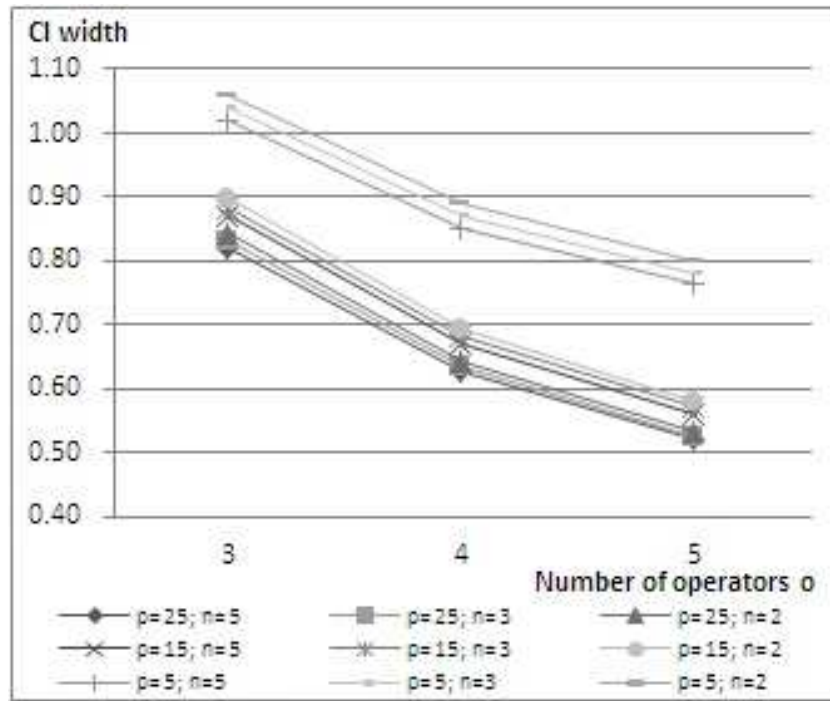
Similar to the previous graphic, Figure 4.38 shows that number of operators has a moderate effect on the width of the confidence interval. As an example, a selected parameter allocation combination (p=25,o=4,n=5) returns an average CI width of 0.63. An increment of 1 operator in the design modifies the average width to 0.52. This change represents a reduction of 17.0%.

Figure 4.39 indicates that there is no significant effect of number of replications on the



* Average values from 625 different combinations of variability components

Figure 4.37: Average CI width of C_p for different sample size p



* Average values from 625 different combinations of variability components

Figure 4.38: Average CI width of C_p for different number of operator o

reduction of the average width of the confidence interval. No matter the number of operators and samples, all scenarios analyzed in the graphic show a small decrement when increasing the number of replications. For example, a Gauge R&R study with parameter allocation combination ($p=5, o=3, n=4$) produces an average CI width of 1.03. One more replication in previous design gives an average width of 1.02, obtaining a reduction of just 0.7%.

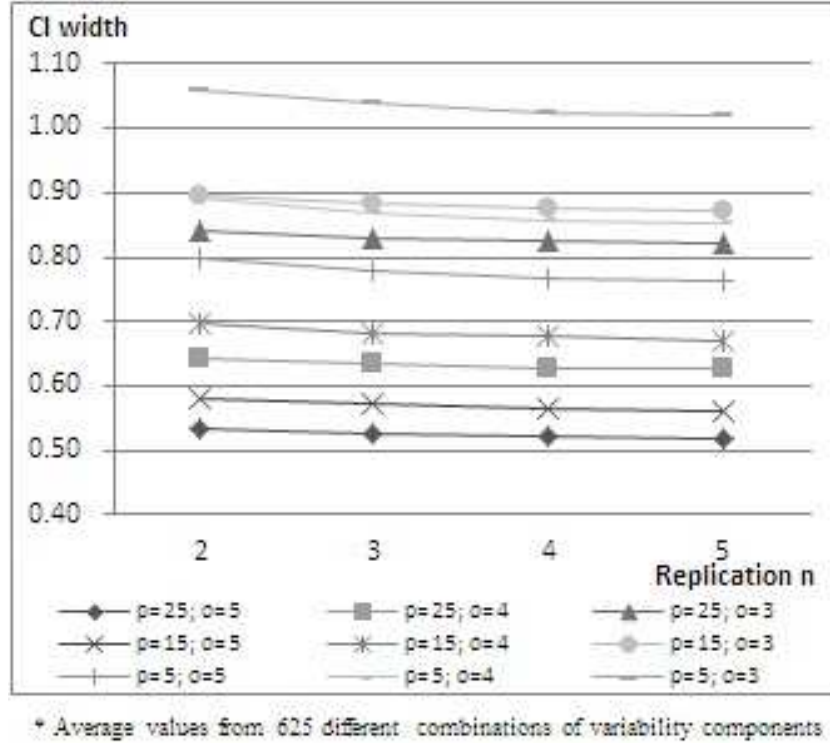


Figure 4.39: Average CI width of C_p for different replication n

Table 4.14 indicates that when having a fixed number of measurements in the Gauge R&R study, it is important to assure a large number of samples and operators in the design for obtaining good results, no matter if the number of replications is minimized. This is clearly evident on the scenario with 300 total measurements and variability components $\sigma_{Rep}^2 = 1$, $\sigma_p^2 = 1$, $\sigma_o^2 = 1$ and $\sigma_{op}^2 = 1$, where the best scenario has the parameter allocation combination ($p=20, o=5, n=3$) that results in an average width of 0.54. Obtaining a reduction of 36.5% with respect to the worst scenario ($p=20, o=3, n=5$) that has an average width of 0.85.

Table 4.14: Average CI width of C_p for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	0.92	0.9	0.97	0.95	0.96	0.96	0.79	0.88	0.93
	10	4	3	0.73	0.71	0.77	0.77	0.78	0.77	0.6	0.69	0.74
	15	4	2	0.69	0.63	0.72	0.76	0.74	0.74	0.58	0.65	0.71
	20	3	2	0.87	0.79	0.89	0.92	0.91	0.91	0.76	0.82	0.88
200	10	4	5	0.72	0.7	0.76	0.75	0.77	0.76	0.56	0.67	0.72
	10	5	4	0.61	0.62	0.65	0.64	0.66	0.65	0.46	0.57	0.61
	20	5	2	0.55	0.48	0.57	0.63	0.61	0.6	0.42	0.5	0.56
	25	4	2	0.64	0.54	0.66	0.73	0.7	0.69	0.51	0.59	0.66
300	15	4	5	0.67	0.62	0.71	0.73	0.72	0.72	0.51	0.62	0.68
	15	5	4	0.56	0.53	0.59	0.62	0.62	0.61	0.4	0.51	0.57
	20	3	5	0.85	0.78	0.88	0.89	0.89	0.89	0.69	0.79	0.86
	20	5	3	0.54	0.48	0.56	0.62	0.6	0.59	0.39	0.49	0.56
	25	3	4	0.83	0.74	0.86	0.89	0.87	0.88	0.68	0.77	0.84
	25	4	3	0.64	0.54	0.66	0.72	0.69	0.69	0.49	0.58	0.65

4.3.2 Analysis of actual potential capability of the process C_p^*

Confidence interval of actual potential capability of the process C_p^* has a behavior similar to the confidence interval of C_p , although in general its average width is twice the size of previous metric. It is shown in Figure 4.40 that the effect of increasing the sample size for reducing the average width is moderate. Considering a design with (p=20,o=4,n=2) it is obtained an average CI width of 1.37, and by increasing 5 samples the new average CI width is 1.20. This represents a reduction of 12.2%.

Figure 4.41 indicates also that the effect of number of operators is moderate in the average confidence interval widths. Similar to the graphic of the confidence interval of repeatability $\sigma_{Repeatability}$, scenarios with small sample size present an abnormal behavior. Rest of scenarios show the expected normal tendency. Six out of nine cases indicates that the effect of increasing the number of operators is moderate on the width of the confidence interval. In order to expose this conclusion it can be considered the design (p=15,o=4,n=2) with an average

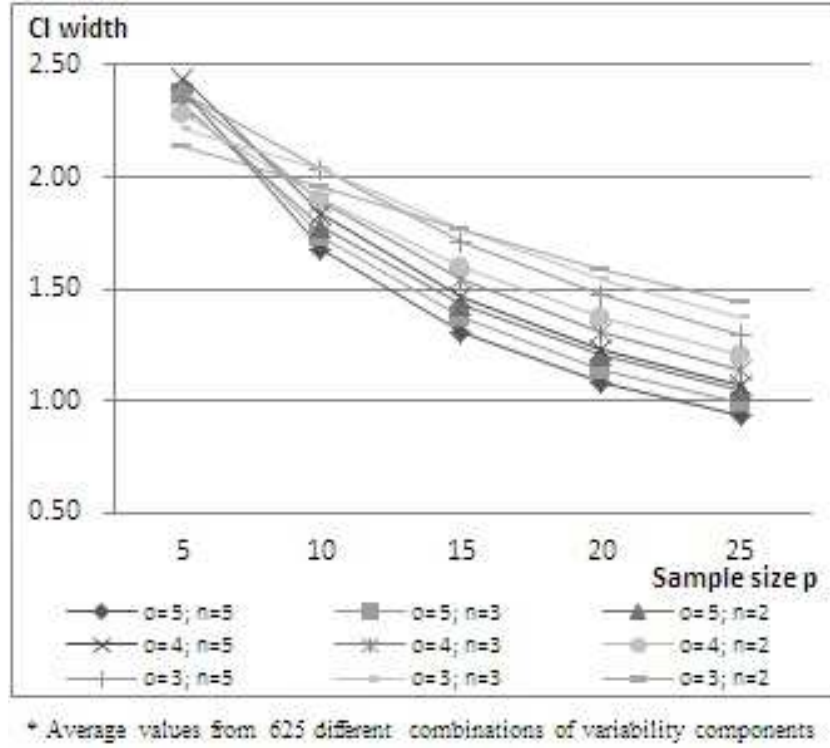


Figure 4.40: Average CI width of C_p^* for different sample size p

width of 1.59. By increasing 1 operator, the average width of the new design ($p=15, o=5, n=2$) changes to 1.43. This represents a reduction of 10.1%.

Figure 4.42 also shows an atypical behavior for three of nine scenarios analyzed. Other six scenarios show almost no reduction of average CI width when increasing the number of replications, meaning that a change in this parameter is not significant. As example, parameter allocation combination ($p=25, o=5, n=3$) has an average width of 0.99. Increasing 1 replication to the design modifies the average width to 0.96, that is a reduction of only 2.1%.

Table 4.15 is used to demonstrate which is the best parameter allocation combination when the total measurements for a Gauge R&R study are limited. In concordance to previous graphics, it can be observed that designs that prioritize sample size and number of operators are better than the ones that include more replications. In the case with a total of 120

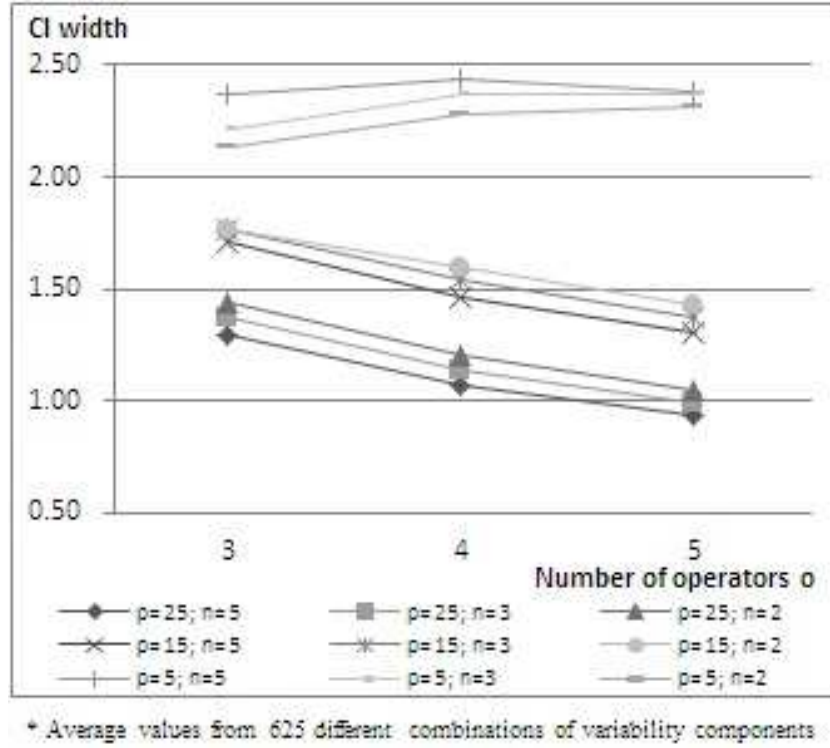


Figure 4.41: Average CI width of C_p^* for different number of operator o

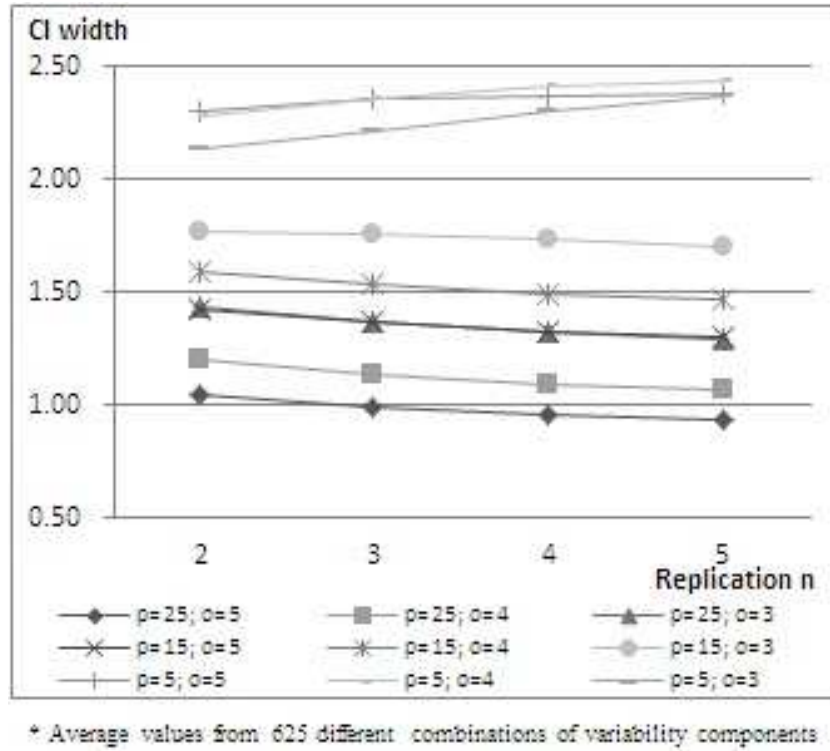


Figure 4.42: Average CI width of C_p^* for different replication n

measurements and variability components $\sigma_{Rep}^2 = 3$, $\sigma_p^2 = 1$, $\sigma_o^2 = 3$ and $\sigma_{op}^2 = 1$, the best design is (p=20,o=3,n=2) which has an average CI width of 1.62. Latter reduces 20.2% the average width compared to the worst scenario.

Table 4.15: Average CI width of C_p^* for different total measurement numbers

			σ_{Rep}^2	1	1	1	3	3	3	5	5	5
			σ_p^2	1	5	3	1	5	3	1	5	3
			σ_o^2	1	1	3	3	5	5	1	3	5
Measurements	p	o	σ_{op}^2 n	1	3	5	1	3	5	1	3	5
120	10	3	4	2.06	1.50	2.32	2.03	1.67	2.41	1.74	1.72	2.45
	10	4	3	1.84	1.31	2.09	2.01	1.45	2.20	1.83	1.53	2.30
	15	4	2	1.49	0.99	1.69	1.69	1.12	1.84	1.76	1.23	2.00
	20	3	2	1.53	0.94	1.75	1.62	1.10	1.94	1.64	1.22	2.07
200	10	4	5	1.76	1.28	2.05	1.98	1.39	2.13	1.97	1.44	2.20
	10	5	4	1.58	1.18	1.83	1.84	1.27	1.94	1.90	1.32	2.01
	20	5	2	1.02	0.76	1.17	1.37	0.81	1.30	1.53	0.87	1.43
	25	4	2	1.01	0.70	1.15	1.40	0.76	1.33	1.52	0.83	1.47
300	15	4	5	1.32	0.96	1.61	1.55	1.01	1.70	1.73	1.06	1.78
	15	5	4	1.16	0.89	1.40	1.41	0.93	1.49	1.58	0.97	1.56
	20	3	5	1.33	0.89	1.69	1.61	0.96	1.79	1.77	1.02	1.83
	20	5	3	0.96	0.75	1.13	1.23	0.79	1.23	1.44	0.82	1.32
	25	3	4	1.15	0.77	1.49	1.48	0.82	1.58	1.67	0.88	1.67
	25	4	3	0.93	0.69	1.14	1.26	0.73	1.24	1.46	0.77	1.36

4.4 Comparison of results

Present research offers a sensitivity analysis of parameter allocation on Gauge R&R study results. Figure 4.43 summarizes the effect of increasing the sample size on the confidence interval average width of the main variability components and capability criteria. From previous findings, it was evidenced that this parameter has a high influence on reducing the average width of two criteria: ρ_p and SNR. For the other seven criteria it was considered that the effect of this parameter on the reduction of the confidence interval width is moderate.

Figure 4.44 presents the influence of the number of operators on the average CI width of the principal criteria under study. It was observed that there exist a high effect of this parameter

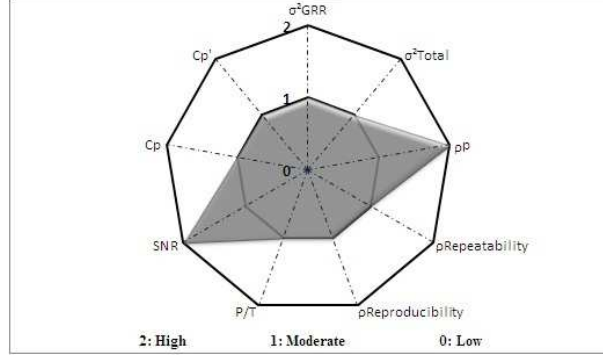


Figure 4.43: Effect of sample size on average confidence interval width

on the CI width of three parameters: σ_{GRR}^2 , σ_{Total}^2 and P/T. For the other six metrics, the effect of this parameter on their respective confidence intervals seemed to be moderate.

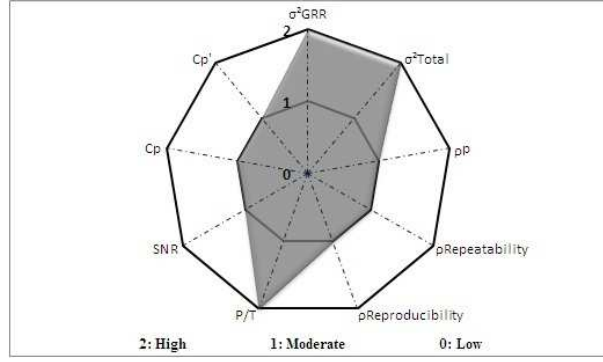


Figure 4.44: Effect of number of operators on average confidence interval width

Following Figure 4.45 indicates the effect of number of replications on the reduction of the average confidence interval width. It was found that this parameter only has a moderate effect on the confidence interval width of two metrics: $\rho_{Repeatability}$ and $\rho_{Reproducibility}$. For all the other metrics the effect of number of replications on the average confidence interval width was considered to be not significant enough.

Considering the results, it is recommended to prioritize first the number of operators in the Gauge R&R study design because it is of benefit for the reduction of the average CI width of a larger number of metrics and variability components. Secondly, it is recommended to increase as much as possible the sample size in the design. This parameter shows a similar benefit

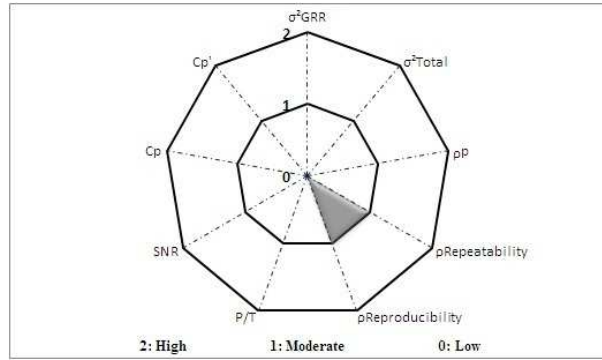


Figure 4.45: Effect of replications on average confidence interval width

in comparison to previous one, and for some metrics results were better by prioritizing an increment on the number of parts. Finally, in general it is clear that there is not a substantial benefit on increasing the number of replications. It is suggested that when possible the level of this parameter should be reduced in order to increase level of previous parameters.

Chapter 5

Case study

On previous chapter, numerical simulation was conducted to perform a sensitivity analysis to establish the effect of parameter allocation combination on the confidence interval width of variability components and various capability criteria used in a Gauge R&R study. A series of recommendations were given to establish the best experimental design depending on the metrics of interest. Previous findings are illustrated on the present chapter with a real case of study.

5.1 General background

Aerospace, pharmaceutical and food industries are among the companies with higher quality requirements. This makes it critical for them to have a reliable measurement system in order to detect anomalies in production process. The following discussion concerns to a real manufacturing process in aerospace industry. Samples were provided by company and the measurement process was replicated for the Gauge R&R study. All data and calculations presented have been modified in order to protect confidentiality policies of the company where the study was performed.

An aircraft is composed of thousands of different parts and subassemblies. One of these parts are the metal plates used in the structure of the fuselage. Each of these plates is braced in the airplane's body employing rivets. Position where rivets are going to be inserted is signaled with a spot in the plate by using a special machine. The operator makes the spot in the surface of the plate with respect to four red marks as shown in Figure 5.1. Before passing the metal plates to the assembly workstation, an off-centering revision of the spot

with respect to the four red marks is performed. Depending on the results the metal piece could be considered as a nonconformity and returned to be reworked.

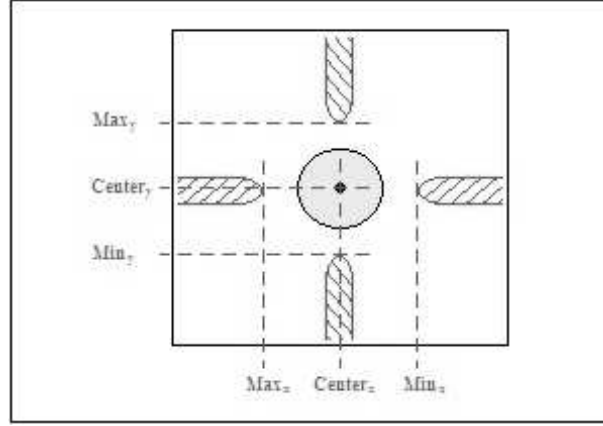


Figure 5.1: Manufactured metal piece

A practitioner is in charge of measuring six values in the metal plate: maximum in axis X Max_x , center in axis X $Center_x$, minimum in axis X Min_x , maximum in axis Y Max_y , center in axis Y $Center_y$ and minimum in axis Y Min_y . Off-centering error is then calculated as

$$Off - center = \sqrt{\left[Center_x - \frac{1}{2}(Max_x + Min_x)\right]^2 + \left[Center_y - \frac{1}{2}(Max_y + Min_y)\right]^2} \quad (5.1)$$

A value of off-center between 0 and 300 is considered acceptable for the assembly process. Any other value will make the metal plate to be rejected.

5.2 Gauge R&R study design

Three capability criteria, ρ_p , P/T and C_p^* , were selected for analyzing reliability of the measurement system and production process. Three different Gauge R&R study designs have been considered in order to verify the effect of number of operators in the results obtained. First experiment consists of 10 samples, 2 operators and 2 replications. Second experiment

includes 10 samples, 3 operators and 2 replications. The last experiment considers 10 samples, 4 operators and 2 replications.

Measurements were performed in random order to assure independency of the observations. Data obtained and the corresponding results of off-centering error calculated are presented in the *Appendix B*.

5.3 Data validation

Before calculating the point estimators of interest and their respective confidence intervals, it is of need to verify the normality assumption of the data collected. Figure 5.2 presents the normal probability plot of the off-centering errors from experiment 1. It is noticed how points fall in the straight line or near it, indicating that the data follows a normal distribution behavior.

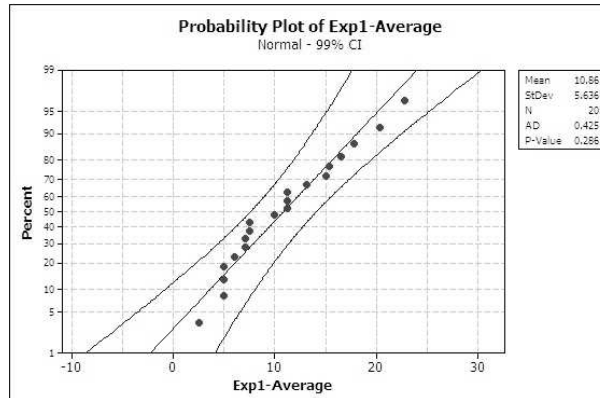


Figure 5.2: Probability plot for normal distribution for experiment 1

Through visual examination of Figure 5.3, it is also verified that all points falls between the confidence interval of the test. Meaning that it can be accepted the normality assumption of data from experiment 2.

Figure 5.4 shows the normal probability plot from experiment 3. It is noticed some points not close to the straight line, but that are still inside the confidence interval of the test. It is

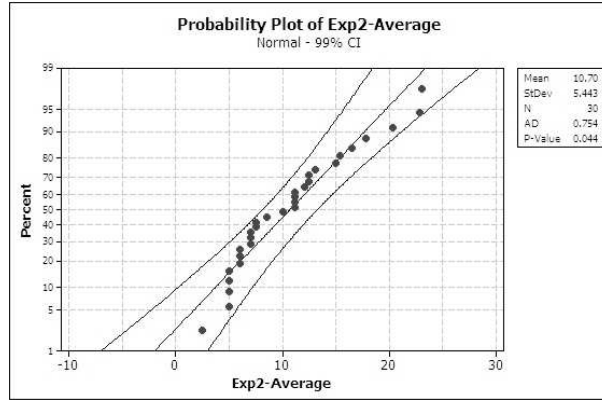


Figure 5.3: Probability plot for normal distribution for experiment 2

considered that for the third experiment normal distribution is an acceptable model for the data.

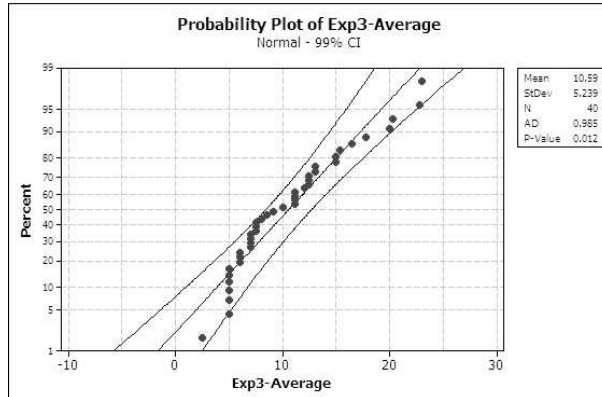


Figure 5.4: Probability plot for normal distribution for experiment 3

It is also of interest to analyze suitability of the measurement instrument to the intended purpose and the capability of the operator to perform the expected work. This can be achieved by interpreting the $\bar{x} - R$ chart. Figure 5.5 presents the $\bar{x} - R$ control chart for experiment 1. As expected, having points out-of-control in the \bar{x} chart is a sign that the measurement instrument is able to differentiate between samples, implying that the measurement instrument is adequate for the activity. On the other hand, R-chart shows an in-control behavior. This indicates that measurements taken by the operators were consistent in each replication. Therefore, it can be considered that the operators did not have difficulties

using the measurement instrument.

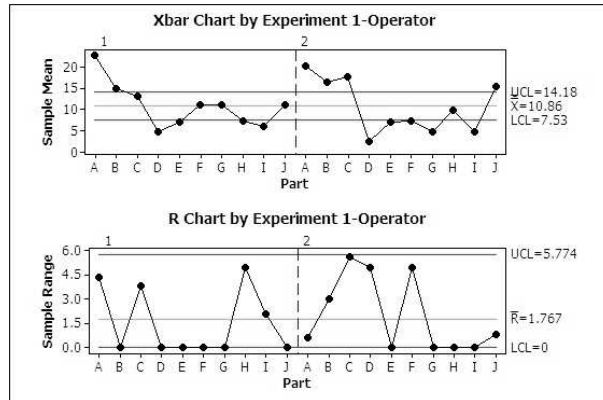


Figure 5.5: $\bar{x} - R$ chart for experiment 1

Similar to previous graphic, it can be noticed on Figure 5.6, that \bar{x} control chart for experiment 2 exhibits several points out-of-control, suggesting that the measurement instrument can distinguish one sample from another. R control chart shows an in-control state. This suggests that there is not a significant difference between the repeated measurements done by the operator on the same sample.

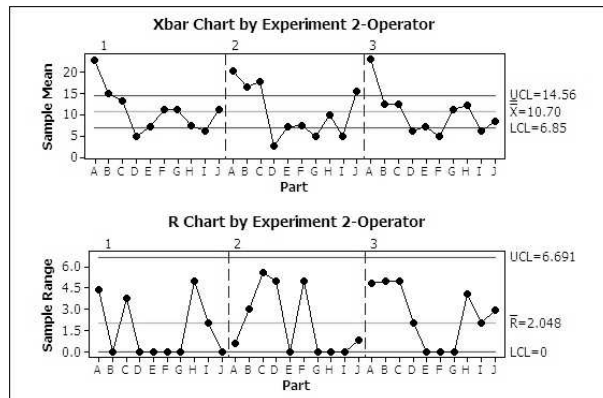


Figure 5.6: $\bar{x} - R$ chart for experiment 2

Figure 5.7 presents results from experiment 3. According to the \bar{x} chart, it is evident again that the gauge is of use for this measurement process. The R chart indicates that the four operators are able to correctly manipulate the measurement instruments and to obtain reliable observations.

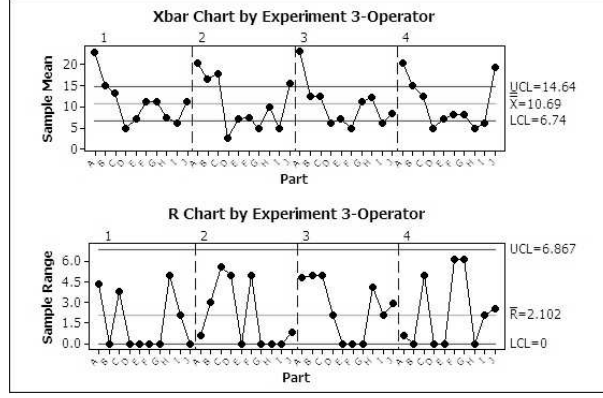


Figure 5.7: $\bar{x} - R$ chart for experiment 3

5.4 Contrast of scenarios

Point estimators and the respective MLS confidence intervals were calculated using analysis of variance method. ANOVA tables for the three experiments are shown in the *Appendix C*. Point estimator of part variability to gauge variability ratio ρ_p for all experiments shows that the variability of the process is more than 3 times the variability of the measurement system, implying that the higher percentage of the error reflected in the measurements is due to the manufacturing process. But due to sampling error, it is also necessary to consider results obtained from the confidence intervals calculated with 5% significance level. For experiment 1, the CI of ρ_p is between 0.31 and 12.96. Having a value of 0.31 implies that most of the variability is due to the gauge, and on the other hand, a value of 12.96 indicates that variability due to gauge is not significant. With this confidence interval it is not possible to conclude about the relation between process and gauge variability. In experiment 2, the width of the confidence interval reduces 9.13% with respect to the point estimator. As expected from previous findings, increasing number of operators has just a moderate effect on reducing the CI width of this metric. The new confidence interval obtained is [1.05, 11.26]. In this case the CI is clearly indicating that at least variability of the gauge is equal or less

to the variability of the process. In last experiment, having a CI equal to $[1.21, 10.51]$ allows to conclude that gauge variability has a lower influence on total error compare to process variability. Although increasing number of operators has not a high impact on reducing the CI width, it is of use to accurately conclude. Table 5.1 summarizes these findings.

Table 5.1: ρ_p results

Experiment	Point estimator	L	U	Width	% reduction width
1	3.57	0.31	12.96	3.55	
2	3.17	1.05	11.26	3.22	9.13%
3	3.00	1.21	10.51	3.10	3.93%

According to AIAG, a measurement system can be considered to be adequate if the precision-to-tolerance ratio is less than 0.1. With respect only to the point estimators of all experiments shown in Table 5.2, the measurement system could be considered as adequate, but again confidence intervals must be considered for a proper discussion. The CI with 5% significance level for experiment 1 is $[0.04, 0.15]$. Having an upper limit greater than 0.1 prevents of concluding about the adequacy of the measurement instrument. By increasing 1 operator in the Gauge R&R study, it is noticed a reduction of 68.81% of the CI width. In experiment 2 all possible values inside the CI of P/T are less than 0.1. It can be concluded that the measurement instrument is adequate for this activity. CI of final experiment is reduced even more, and it is more evident the adequacy of the gauge.

Table 5.2: P/T results

Experiment	Point estimator	L	U	Width	% reduction width
1	0.06	0.04	0.15	1.87	
2	0.06	0.05	0.08	0.58	68.81%
3	0.06	0.05	0.07	0.43	25.55%

Table 5.3 presents actual process capability ratio C_p^* . A process with a value of this metric greater than 1.3 is considered to be capable. For all experiments, point estimators and confidence intervals are greater than 1.3. For this metric it was of no need to increase the number of operator to conclude that the process is capable. It has to be noticed also that increasing the number of operators has only a moderate effect on reducing the CIŠs width of this metric.

Table 5.3: C_p^* results

Experiment	Point estimator	L	U	Width	% reduction width
1	9.58	5.05	15.68	1.11	
2	9.92	5.29	15.27	1.01	9.33%
3	9.91	5.31	15.02	0.98	2.55%

This example problem illustrates that depending on how the Gauge R&R study is designed, conclusions from the capability metrics can be more or less significant. In this case, increasing number of operators helped to conclude about ρ_p and P/T.

Chapter 6

Conclusions

6.1 Summary

In this thesis, effects of Gauge R&R study parameter allocation on the MLS confidence intervals of variability components and capability criteria were studied through numerical simulation. Based on the literature review of the research on measurement system analysis in Chapter 2, it was introduced the point estimators and confidence intervals of interest in Chapter 3. Findings of the simulation were summarized in Chapter 4. A numerical example was given in Chapter 5 to illustrate the analysis performed.

6.2 Contributions of the thesis

A reliable measurement system is of need for inspecting and improving quality of the process. Several indices are used in industry to evaluate adequacy of the measurement system. Gauge R&R study is among the existing tools used to calculate capability indices point estimates in measurement system analysis. But because of sampling uncertainty, it is required estimation of confidence intervals for the analysis. This thesis focused on the effect of Gauge R&R study design on the calculation of the confidence intervals of variability components and capability indices using MLS method.

Simulation results on the present research lead to conclude that:

- Sample size has a high impact on the width of the confidence interval for SNR and ρ_p , and a moderate impact for σ_{GRR}^2 , σ_{Total}^2 , $\rho_{Repeatability}$, $\rho_{Reproducibility}$, P/T, C_p and C_p^* .
- Number of operators has major effect on the width of the confidence interval for σ_{GRR}^2 , σ_{Total}^2 and P/T, and a moderate effect for ρ_p , $\rho_{Repeatability}$, $\rho_{Reproducibility}$, SNR, C_p and C_p^* .

- Number of replications has a moderate effect on the width of the confidence interval for $\rho_{Repeatability}$ and $\rho_{Reproducibility}$, and a low effect for P/T, SNR, C_p and C_p^* , σ_{GRR}^2 , σ_{Total}^2 and ρ_p .

Based on these results, it is recommended to prioritize number of operators and sample size in the design of a Gauge R&R study. This will be of benefit for the reduction of a larger number of metrics and variability components. Number of replications has no a significant effect on the reduction of CI width. Therefore, it is suggested that when possible the number of replications have to be reduced in order to increase number of latter parameters.

6.3 Future research

Present research focused on modified-large-sample method to estimate CIs of variability components and capability criteria. Further study on how Gauge R&R study design affects estimations of others methods, like Satterthwaite (SATT) or maximum likelihood (RELM) method, is of interest. Moreover, conclusions drawn are restricted to particular scenarios considered as the most common in industry. The present research can be extended by considering the effect of distinct parameter allocations on Gauge R&R study. Capability of the measurement system can also be studied by analyzing its ability to correctly differentiate between conformities and nonconformities (Woodall and Borror, 2008). Therefore, it will be also important to include misclassification rates on the sensitivity analysis. Gauge R&R experimental design not only influences reliability of the results. Depending on the personnel, parts and replications selected for the study, the time and cost required can highly varied. Including both variables in the analysis will complement the findings of this research.

References

1. Al-Refaie, A. (2011), "Evaluating measurement and process capabilities using tabular algorithm procedure with three quality measures", *Transactions of the Institute of Measurement and Control*, Vol. 34, pp. 604–614.
2. Al-Refaie, A. and Bata, N. (2010), "Evaluating measurement and process capabilities by GR&R with four quality measures", *Measurement: Journal of the International Measurement Confederation*, Vol. 43, pp. 842–851.
3. Antony, J., Knowles, G. and Roberts, P. (1999), "Gauge capability analysis: classical versus ANOVA", *Quality Assurance: Good Practice, Regulation, and Law*, Vol. 6, pp. 173–181.
4. Arteaga, C., Jeyaratnam, S. and Graybill, F. (1982), "Confidence intervals for proportions of total variance in the two-way cross component of variance model", *Communications in Statistics-Theory and Methods*, Vol. 11, pp. 1643–1658.
5. Automotive Industry Action Group AIAG (2002), "Measurement Systems Analysis Reference Manual", Detroit, Michigan.
6. Bordignon, S. and Scagliarini, M. (2002), "Statistical analysis of process capability indices with measurement errors", *Quality and reliability engineering international*, Vol. 18, pp. 321–332.
7. Borror, C., Montgomery, D. and Runger, G. (1997), "Confidence intervals for variance components from gauge capability studies", *Quality and Reliability Engineering International*, Vol. 13, pp. 361–369.

8. Burdick, R. and Larsen, G. (1997), "Confidence intervals on measures of variability in R&R studies", *Journal of Quality Technology*, Vol. 29, pp. 261–273.
9. Burdick, R., Borror, C. and Montgomery, D. (2003), "A review of methods for measurement systems capability analysis", *Journal of Quality Technology*, Vol. 35, pp. 342–354.
10. Burdick, R., Borror, C. and Montgomery, D. (2005), "Design and analysis of gauge R&R studies: Making decisions with confidence intervals in random and mixed ANOVA models", SIAM.
11. Dasgupta, T. and Murthy, S. (2001), "Looking beyond audit-oriented evaluation of gauge repeatability and reproducibility: A case study", *Total Quality Management*, Vol. 12, pp. 649–655.
12. Dejaegher, B., Jimidar, M., De Smet, M., Cockaerts, P., Smeyers-Verbeke, J. and Vander Heyden, Y. (2006), "Improving method capability of a drug substance HPLC assay", *Journal of pharmaceutical and biomedical analysis*, Vol. 42, pp. 155–170.
13. Gao, Z., Moore, T., Smith, A., Doub, W., Westenberger, B. and Buhse, L. (2007), "Gauge repeatability and reproducibility for accessing variability during dissolution testing: a Technical Note", *AAPS PharmSciTech*, Vol. 8, pp. 11–15.
14. Graybill, F. and Wang, C. (1980), "Confidence intervals on nonnegative linear combinations of variances", *Journal of the American Statistical Association*, Vol. 75, pp. 869–873.
15. Harry, M. and Lawson, J. (1992), "Six sigma producibility analysis and process characterization", Addison-Wesley.

16. He, Z., Liu, B. and Qi, E. (2000), "A study on measurement system analysis for different measurement conditions", Proceedings of the Seventh International Conference on Industrial Engineering and Engineering Management, 28–30 November, pp. 408–412, Guangzhou, China.
17. Houf, R. and Berman, D. (1988), "Statistical analysis of power module thermal test equipment performance", *Components, Hybrids, and Manufacturing Technology, IEEE Transactions on*, Vol. 11, pp. 516–520.
18. Hsu, B., Shu, M. and Pearn, W. (2007), "Measuring process capability based on Cpmk with gauge measurement errors", *Quality and Reliability Engineering International*, Vol. 23, pp. 597–614.
19. Ile, C. and Lazea, G. (2014), "Use Measurement System Analysis in Order to Reduce the System Variation", *Applied Mechanics and Materials*, Vol. 436, pp. 295–302.
20. Ingram, D. and Taylor, W. (1998), "Measurement system analysis", Proceedings of the ASQ World Conference on Quality and Improvement Proceedings, 4–6 May, pp. 931–941, Philadelphia, USA.
21. Kaija, K., Pekkanen, V., Mantysalo, M., Koskinen, S., Niittynen, J., Halonen, E. and Mansikkamaki, P. (2010), "Inkjetting dielectric layer for electronic applications", *Micro-electronic Engineering*, Vol. 87, pp. 1984–1991.
22. Larsen, G. (2003), "Measurement system analysis in a production environment with multiple test parameters", *Quality Engineering*, Vol. 16, pp. 297–306.

23. Leiva, R. and Graybill, F. (1986), "Confidence intervals for variance components in the balanced two way model with interaction", *Communications in Statistics-Simulation and Computation*, Vol. 15, pp. 301–322.
24. Li, M. and Al-Refaie, A. (2008), "Improving wooden parts' quality by adopting DMAIC procedure", *Quality and Reliability Engineering International*, Vol. 24, pp. 351–360.
25. Lin, C., Hong, C. and Lai, J. (1997), "Improvement of a dimensional measurement process using Taguchi robust designs", *Quality Engineering*, Vol. 9, pp. 561–573.
26. Lu, T., Graybill, F. and Burdick, R. (1987), "Confidence Intervals on the Ratio of Expected Mean Squares $(\theta_1 + d\theta_2)/\theta_3$ ", *Biometrics*, Vol. 43, pp. 535–543.
27. Mader, D., Prins, J. and Lampe, R. (1999), "The economic impact of measurement error", *Quality Engineering*, Vol. 11, pp. 563–574.
28. Majeske, K. and Andrews, R. (2002), "Evaluating measurement systems and manufacturing processes using three quality measures", *Quality Engineering*, Vol. 15, pp. 243–251.
29. Montgomery, D. (2012), "Introduction to statistical quality control", Wiley Global Education.
30. Montgomery, D. and Runger, G. (1993), "Gauge capability analysis and designed experiments. Part II: experimental design models and variance component estimation", *Quality Engineering*, Vol. 6, pp. 289–305.
31. Montgomery, D. and Runger, G. (1993), "Gauge capability and designed experiments. Part I: basic methods", *Quality Engineering*, Vol. 6, pp. 115–135.

32. Morris, R. and Watson, E. (1998), "A comparison of the techniques used to evaluate the measurement process", *Quality Engineering*, Vol. 11, pp. 213–219.
33. Osma, A. (2011), "An assessment of the robustness of gauge repeatability and reproducibility analysis in automotive components", Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, Vol. 225, pp. 895–912.
34. Pan, J. (2004), "Determination of the optimal allocation of parameters for gauge repeatability and reproducibility study", *International Journal of Quality & Reliability Management*, Vol. 21, pp. 672–682.
35. Pan, J. (2006), "Evaluating the gauge repeatability and reproducibility for different industries", *Quality and Quantity*, Vol. 40, pp. 499–518.
36. Park, D. and Yoon, M. (2009), "Optimal confidence intervals on the variation for operators and gauge in manufacturing process with repeated measurements", *Optimization and Engineering*, Vol. 10, pp. 273–287.
37. Pearn, W. and Liao, M. (2005), "Measuring process capability based on Cpk with gauge measurement errors", *Microelectronics Reliability*, Vol. 45, pp. 739–751.
38. Pearn, W. and Liao, M. (2007), "Estimating and testing process precision with presence of gauge measurement errors", *Quality and Quantity*, Vol. 41, pp. 757–777.
39. Peruchi, R., Balestrassi, P., de Paiva, A., Ferreira, J. and de Santana, M. (2013), "A new multivariate gage R&R method for correlated characteristics", *International Journal of Production Economics*, Vol. 144, pp. 301–315.

40. Raffaldi, J. and Kappele, W. (2004), "Improve Gage R & R Ratios", *Quality*, Vol. 43, pp. 48–54.
41. Ramesh, B. and Sarma, K. (2013), "On the estimation of variance components in Gage R&R studies using Ranges & ANOVA", *International Journal of Scientific & Engineering Research*, Vol. 4, pp. 1900–1905.
42. Senvar, O. and Firat, S. (2010), "Investigation of gauge and measurement systems capability studies", Proceedings of the ENBIS - IMEKO TC21 Workshop on Measurement Systems and Process Improvement, 19–20 April, pp. 100–109, Teddington, United Kingdom.
43. Senvar, O. and Oktay, S. (2010), "An overview of capability evaluation of Measurement Systems and Gauge Repeatability and Reproducibility Studies", *International Journal of Metrology and Quality Engineering*, Vol. 1, pp. 121–127.
44. Sivaji, A. (2006), "Measurement system analysis", Proceedings of the Third IEEE International Workshop on Electronic Design, Test and Applications, 17–19 January, pp. 393–396, Kuala Lumpur, Malaysia.
45. Smith, R., McCrary, S. and Callahan, R. (2007), "Gauge repeatability and reproducibility studies and measurement system analysis: a multimethod exploration of the state of practice", *Journal of Industrial Technology*, Vol. 23, pp. 1–11.
46. Srikaeo, K., Furst, J. and Ashton, J. (2005), "Characterization of wheat-based biscuit cooking process by statistical process control techniques", *Food Control*, Vol. 16, pp. 309–317.

47. Ting, N., Burdick, R., Graybill, F., Jeyaratnam, S. and Lu, T. (1990), "Confidence intervals on linear combinations of variance components that are unrestricted in sign", *Journal of Statistical Computation and Simulation*, Vol. 35, pp. 135–143.
48. Tsai, P. (1988), "Variable gauge repeatability and reproducibility study using the analysis of variance method", *Quality Engineering*, Vol. 1, pp. 107–115.
49. Wang, F. (2008), "Process yield with measurement errors in semiconductor manufacturing", *Semiconductor Manufacturing, IEEE Transactions on*, Vol. 21, pp. 279–284.
50. Wang, F. and Li, E. (2003), "Confidence intervals in repeatability and reproducibility using the Bootstrap method", *Total Quality Management and Business Excellence*, Vol. 14, pp. 341–354.
51. Wheeler, D. (2009), "An honest gauge R&R study", 1–19, <http://www.spcpress.com/pdf/DJW189.pdf>.
52. Wheeler, D. and Lyday, R. (1989), "Evaluating the measurement process", SPC Press.
53. Woodall, W. and Borror, C. (2008), "Some relationships between gage R&R criteria", *Quality and Reliability Engineering International*, Vol. 24, pp. 99–106.
54. Zappa, D. and Deldossi, L. (2009), "Misclassification rates, critical values and size of the design in measurement systems capability studies", *Applied Stochastic Models in Business and Industry*, Vol. 25, pp. 601–611.

Appendices

A Numerical simulation data example

Table A.1: Simulation data example for average confidence interval width of σ_{Rep}^2

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	1.49	1.66	1.75	1.78	1.73	1.77	1.43	1.49	1.51	1.39	1.39	1.44	1.48	
5	3	3	1.08	1.13	1.14	1.14	1.14	1.14	1.05	1.08	1.08	1.02	1.01	1.05	1.08	
5	3	4	0.88	0.89	0.90	0.90	0.90	0.90	0.86	0.88	0.88	0.84	0.83	0.86	0.88	
5	3	5	0.75	0.76	0.76	0.76	0.76	0.76	0.75	0.75	0.75	0.73	0.73	0.75	0.75	
5	4	2	1.29	1.45	1.49	1.49	1.47	1.48	1.25	1.30	1.31	1.17	1.14	1.22	1.30	
5	4	3	0.94	0.96	0.96	0.96	0.96	0.96	0.92	0.94	0.94	0.88	0.86	0.91	0.94	
5	4	4	0.76	0.76	0.76	0.76	0.76	0.76	0.75	0.76	0.76	0.73	0.71	0.75	0.76	
5	5	2	1.17	1.28	1.29	1.29	1.28	1.29	1.14	1.17	1.18	1.05	1.00	1.10	1.17	
5	5	3	0.83	0.84	0.85	0.85	0.85	0.85	0.83	0.83	0.84	0.80	0.76	0.81	0.83	
5	5	4	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.66	0.63	0.67	0.67	
5	5	5	0.58	0.58	0.58	0.58	0.58	0.58	0.57	0.58	0.58	0.57	0.56	0.57	0.58	
10	3	2	1.07	1.14	1.15	1.15	1.15	1.15	1.03	1.07	1.08	0.94	0.91	1.00	1.07	
10	3	3	0.76	0.76	0.76	0.76	0.76	0.76	0.75	0.76	0.76	0.72	0.69	0.74	0.76	
10	3	4	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.60	0.58	0.61	0.61	
10	3	5	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.51	0.52	0.52	
10	4	2	0.94	0.96	0.96	0.96	0.96	0.96	0.92	0.94	0.94	0.84	0.77	0.88	0.94	
10	4	4	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.50	0.52	0.52	
10	4	5	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.44	0.45	0.45	
10	5	2	0.84	0.85	0.85	0.85	0.85	0.85	0.83	0.84	0.84	0.78	0.69	0.80	0.84	
10	5	3	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.57	0.54	0.57	0.58	
10	5	4	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.45	0.46	0.46	
10	5	5	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.39	0.40	0.40	
15	3	2	0.88	0.90	0.90	0.90	0.90	0.90	0.86	0.88	0.88	0.79	0.73	0.84	0.88	
15	3	3	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.60	0.57	0.60	0.61	
15	3	4	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.47	0.49	0.49	
15	3	5	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	
15	4	2	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.72	0.63	0.73	0.76	
15	4	4	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.41	0.42	0.42	
15	4	5	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	
15	5	2	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.65	0.57	0.66	
15	5	3	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.44	0.46	0.46	
15	5	4	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.37	0.38	0.38	
15	5	5	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	
20	3	3	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.49	0.52	0.52	
20	3	4	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.41	0.42	0.42	
20	3	5	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	
20	4	2	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.63	0.55	0.64	0.65	
20	4	3	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.43	0.45	0.45	
20	4	4	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	
20	5	2	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.57	0.50	0.57	0.58	
25	4	3	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.39	0.40	0.40	
25	4	4	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	
25	4	5	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	
25	5	2	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.45	0.51	0.51	
25	5	3	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.35	0.36	0.36	
25	5	4	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	
25	5	5	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	

Table A.2: Simulation data example for average confidence interval width of σ_p^2

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	9.22	9.39	9.52	9.32	8.96	9.13	9.54	9.23	9.03	9.95	9.28	9.19	9.22	
5	3	3	9.23	9.51	9.59	9.35	8.99	9.14	9.42	9.22	9.07	9.61	9.01	9.09	9.22	
5	3	4	9.28	9.54	9.6	9.36	9.01	9.15	9.48	9.28	9.1	9.48	8.88	9.11	9.29	
5	3	5	9.3	9.55	9.62	9.36	9	9.14	9.53	9.31	9.12	9.49	8.81	9.13	9.31	
5	4	2	9.44	9.86	9.99	9.6	9.08	9.23	9.79	9.44	9.2	9.96	9.19	9.29	9.44	
5	4	3	9.51	9.9	10.03	9.59	9.07	9.21	9.9	9.5	9.23	9.83	8.96	9.27	9.5	
5	4	4	9.53	9.9	10	9.6	9.06	9.19	9.96	9.53	9.23	9.89	8.86	9.28	9.52	
5	5	2	9.55	10.08	10.21	9.67	9.06	9.19	10.05	9.56	9.24	10.07	9.11	9.32	9.56	
5	5	3	9.58	10.07	10.22	9.67	9.04	9.15	10.15	9.58	9.23	10.08	8.89	9.31	9.58	
5	5	4	9.57	10.06	10.2	9.65	9.03	9.14	10.18	9.56	9.2	10.16	8.8	9.29	9.56	
5	5	5	9.54	10.05	10.22	9.64	9.02	9.13	10.18	9.55	9.18	10.21	8.75	9.27	9.54	
10	3	2	3.89	4.18	4.25	3.98	3.58	3.63	4.09	3.89	3.69	4.07	3.54	3.72	3.88	
10	3	3	3.91	4.17	4.25	3.98	3.57	3.61	4.19	3.9	3.68	4.11	3.42	3.73	3.91	
10	3	4	3.88	4.17	4.23	3.97	3.56	3.6	4.21	3.89	3.65	4.19	3.37	3.71	3.89	
10	3	5	3.87	4.16	4.24	3.97	3.55	3.59	4.21	3.87	3.64	4.21	3.34	3.69	3.88	
10	4	2	3.9	4.26	4.37	3.98	3.48	3.51	4.3	3.9	3.62	4.25	3.44	3.7	3.9	
10	4	4	3.83	4.24	4.36	3.96	3.44	3.47	4.31	3.83	3.54	4.34	3.27	3.61	3.83	
10	4	5	3.81	4.24	4.37	3.96	3.44	3.46	4.3	3.8	3.52	4.33	3.24	3.58	3.8	
10	5	2	3.83	4.27	4.41	3.91	3.38	3.4	4.36	3.83	3.52	4.35	3.36	3.62	3.83	
10	5	3	3.76	4.24	4.41	3.89	3.36	3.38	4.34	3.77	3.47	4.38	3.25	3.55	3.76	
10	5	4	3.73	4.23	4.4	3.89	3.35	3.37	4.31	3.72	3.43	4.36	3.2	3.5	3.72	
10	5	5	3.7	4.22	4.4	3.88	3.34	3.36	4.3	3.7	3.41	4.34	3.17	3.47	3.7	
15	3	2	2.9	3.2	3.28	2.98	2.51	2.53	3.22	2.9	2.65	3.13	2.46	2.72	2.9	
15	3	3	2.86	3.19	3.28	2.96	2.49	2.51	3.24	2.86	2.6	3.23	2.36	2.67	2.86	
15	3	4	2.83	3.19	3.28	2.96	2.48	2.49	3.24	2.83	2.56	3.25	2.31	2.63	2.83	
15	3	5	2.81	3.18	3.28	2.96	2.47	2.49	3.22	2.81	2.54	3.25	2.28	2.6	2.81	
15	4	2	2.8	3.21	3.35	2.89	2.38	2.39	3.3	2.8	2.51	3.3	2.35	2.61	2.8	
15	4	4	2.69	3.19	3.33	2.87	2.35	2.36	3.25	2.69	2.42	3.3	2.22	2.48	2.69	
15	4	5	2.67	3.18	3.33	2.86	2.34	2.35	3.24	2.67	2.4	3.28	2.19	2.45	2.67	
15	5	2	2.67	3.16	3.35	2.78	2.29	2.3	3.28	2.68	2.4	3.33	2.27	2.49	2.68	
15	5	3	2.61	3.14	3.34	2.76	2.28	2.28	3.24	2.61	2.35	3.3	2.2	2.42	2.61	
15	5	4	2.57	3.13	3.33	2.75	2.27	2.27	3.21	2.57	2.32	3.26	2.16	2.38	2.57	
15	5	5	2.55	3.12	3.32	2.75	2.26	2.27	3.19	2.55	2.31	3.24	2.14	2.35	2.55	
20	3	3	2.36	2.76	2.86	2.5	1.98	2	2.81	2.36	2.08	2.84	1.88	2.16	2.36	
20	3	4	2.33	2.75	2.86	2.49	1.97	1.98	2.8	2.33	2.05	2.83	1.83	2.11	2.33	
20	3	5	2.3	2.75	2.86	2.48	1.96	1.98	2.79	2.3	2.03	2.82	1.8	2.08	2.3	
20	4	2	2.27	2.74	2.9	2.38	1.88	1.89	2.84	2.27	1.99	2.88	1.87	2.09	2.27	
20	4	3	2.2	2.71	2.89	2.36	1.87	1.87	2.8	2.2	1.94	2.86	1.79	2.01	2.2	
20	4	4	2.16	2.7	2.88	2.35	1.86	1.86	2.78	2.16	1.92	2.83	1.76	1.97	2.16	
20	5	2	2.13	2.65	2.87	2.25	1.81	1.82	2.79	2.14	1.89	2.87	1.8	1.98	2.13	
20	5	3	2.07	2.62	2.85	2.23	1.8	1.8	2.73	2.07	1.85	2.81	1.74	1.91	2.07	
20	5	4	2.04	2.61	2.85	2.22	1.79	1.8	2.7	2.04	1.83	2.76	1.72	1.88	2.04	
20	5	5	2.02	2.6	2.85	2.21	1.79	1.79	2.68	2.02	1.82	2.73	1.7	1.86	2.02	
25	3	2	2.12	2.51	2.63	2.22	1.7	1.71	2.59	2.12	1.82	2.58	1.68	1.93	2.12	
25	3	3	2.05	2.5	2.63	2.2	1.68	1.69	2.56	2.05	1.77	2.6	1.6	1.85	2.05	
25	3	4	2.01	2.49	2.62	2.19	1.67	1.68	2.54	2.01	1.74	2.58	1.56	1.8	2.01	
25	3	5	1.99	2.48	2.62	2.19	1.67	1.67	2.53	1.99	1.72	2.57	1.53	1.77	1.99	
25	4	3	1.87	2.42	2.62	2.03	1.58	1.59	2.51	1.87	1.64	2.58	1.52	1.7	1.87	
25	4	4	1.84	2.4	2.62	2.03	1.58	1.58	2.48	1.84	1.62	2.54	1.5	1.66	1.84	
25	4	5	1.81	2.4	2.61	2.02	1.57	1.57	2.47	1.82	1.61	2.52	1.48	1.64	1.82	
25	5	2	1.81	2.34	2.58	1.92	1.54	1.54	2.48	1.81	1.6	2.58	1.53	1.67	1.81	
25	5	3	1.75	2.31	2.56	1.9	1.53	1.53	2.42	1.75	1.57	2.5	1.48	1.62	1.75	
25	5	4	1.72	2.29	2.56	1.89	1.52	1.52	2.38	1.72	1.56	2.45	1.46	1.59	1.72	
25	5	5	1.71	2.28	2.55	1.88	1.52	1.52	2.36	1.71	1.55	2.42	1.44	1.57	1.71	

Table A.3: Simulation data example for average confidence interval width of σ_o^2

			σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
p	o	n	σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	43.7	44.33	44.79	45.29	45.05	43.2	43.57	43.7	43.83	43.73	43.9	43.68	43.69	
5	3	3	43.62	44.64	45.13	45.47	45.32	43.22	43.07	43.6	43.84	43.47	43.09	43.23	43.6	
5	3	4	43.76	44.75	45.05	45.5	45.35	43.22	43.2	43.73	43.94	42.57	42.58	43.19	43.73	
5	3	5	43.8	44.83	45.1	45.52	45.41	43.24	43.25	43.81	43.99	42.53	42.34	43.24	43.78	
5	4	2	15.79	16.42	16.63	16.81	16.65	15.48	15.6	15.8	15.86	15.43	15.45	15.58	15.79	
5	4	3	15.87	16.48	16.67	16.84	16.68	15.46	15.66	15.88	15.95	15.25	15.13	15.54	15.87	
5	4	4	15.88	16.46	16.64	16.81	16.66	15.43	15.69	15.91	15.95	15.26	14.98	15.54	15.9	
5	5	2	9.55	10.07	10.23	10.31	10.19	9.19	9.42	9.55	9.59	9.18	9.1	9.32	9.56	
5	5	3	9.58	10.07	10.21	10.31	10.19	9.16	9.46	9.58	9.62	9.13	8.89	9.31	9.59	
5	5	4	9.56	10.05	10.22	10.31	10.17	9.14	9.47	9.56	9.59	9.14	8.81	9.29	9.56	
5	5	5	9.55	10.04	10.21	10.29	10.15	9.13	9.44	9.55	9.58	9.13	8.75	9.27	9.54	
10	3	2	44.31	46.02	46.75	47.2	46.62	43.17	43.79	44.3	44.52	43.03	42.91	43.61	44.33	
10	3	3	44.34	45.94	46.65	47.11	46.56	43.08	43.84	44.31	44.49	42.87	42.38	43.55	44.33	
10	3	4	44.23	45.92	46.75	47.13	46.5	43.04	43.84	44.27	44.39	42.93	42.11	43.49	44.27	
10	3	5	44.18	45.93	46.62	47.14	46.49	43.01	43.8	44.17	44.33	42.9	42.01	43.39	44.2	
10	4	2	15.84	16.73	17.09	17.26	16.9	15.1	15.69	15.85	15.89	15.23	14.99	15.45	15.84	
10	4	4	15.7	16.66	17.08	17.22	16.81	15.04	15.57	15.69	15.72	15.14	14.65	15.29	15.69	
10	4	5	15.65	16.64	17.07	17.21	16.79	15.02	15.53	15.65	15.69	15.09	14.59	15.24	15.65	
10	5	2	9.39	10.08	10.39	10.47	10.17	8.8	9.31	9.39	9.4	9	8.73	9.1	9.39	
10	5	3	9.29	10.03	10.36	10.46	10.11	8.77	9.23	9.29	9.31	8.93	8.58	9	9.3	
10	5	4	9.24	10.03	10.36	10.45	10.09	8.76	9.18	9.25	9.25	8.87	8.51	8.94	9.24	
10	5	5	9.21	10	10.35	10.43	10.09	8.74	9.15	9.21	9.21	8.83	8.46	8.9	9.21	
15	3	2	44.08	45.89	46.76	47.2	46.29	42.63	43.68	44.08	44.16	42.77	42.37	43.28	44.04	
15	3	3	43.89	45.76	46.72	47.12	46.22	42.51	43.54	43.91	43.95	42.74	41.9	43.11	43.89	
15	3	4	43.78	45.71	46.66	47.1	46.13	42.47	43.45	43.76	43.83	42.61	41.7	42.95	43.77	
15	3	5	43.7	45.65	46.66	47.08	46.09	42.44	43.39	43.67	43.72	42.51	41.55	42.87	43.68	
15	4	2	15.53	16.49	16.96	17.1	16.6	14.78	15.41	15.54	15.55	15.04	14.7	15.17	15.52	
15	4	4	15.34	16.39	16.93	17.04	16.49	14.71	15.24	15.33	15.35	14.87	14.41	14.95	15.33	
15	4	5	15.29	16.38	16.9	17.05	16.47	14.7	15.22	15.3	15.3	14.81	14.35	14.89	15.3	
15	5	2	9.08	9.83	10.21	10.27	9.88	8.56	9.05	9.09	9.09	8.79	8.52	8.84	9.08	
15	5	3	8.99	9.79	10.18	10.26	9.82	8.53	8.97	9	9	8.69	8.4	8.73	9	
15	5	4	8.95	9.77	10.18	10.23	9.8	8.52	8.92	8.95	8.95	8.65	8.34	8.68	8.94	
15	5	5	8.92	9.74	10.17	10.23	9.79	8.51	8.88	8.91	8.92	8.61	8.3	8.64	8.91	
20	3	3	43.38	45.39	46.44	46.78	45.71	42.07	43.13	43.43	43.47	42.39	41.57	42.69	43.42	
20	3	4	43.26	45.35	46.38	46.77	45.64	42.02	43.04	43.25	43.3	42.26	41.35	42.51	43.25	
20	3	5	43.17	45.32	46.34	46.74	45.56	42	42.96	43.18	43.22	42.16	41.24	42.4	43.18	
20	4	2	15.26	16.22	16.76	16.85	16.28	14.56	15.19	15.26	15.25	14.85	14.5	14.92	15.25	
20	4	3	15.13	16.17	16.71	16.82	16.22	14.52	15.07	15.12	15.13	14.72	14.34	14.79	15.12	
20	4	4	15.06	16.13	16.69	16.81	16.19	14.5	15.01	15.07	15.08	14.66	14.25	14.71	15.07	
20	5	2	8.87	9.6	10.01	10.06	9.62	8.42	8.85	8.88	8.88	8.63	8.39	8.66	8.87	
20	5	3	8.79	9.56	9.99	10.03	9.58	8.39	8.76	8.79	8.79	8.55	8.29	8.57	8.79	
20	5	4	8.75	9.54	9.97	10.03	9.56	8.38	8.73	8.74	8.75	8.49	8.24	8.51	8.75	
20	5	5	8.72	9.52	9.97	10.01	9.54	8.38	8.7	8.72	8.72	8.47	8.21	8.48	8.72	
25	3	2	43.27	45.1	46.13	46.47	45.37	41.82	43.02	43.24	43.26	42.28	41.65	42.58	43.24	
25	3	3	43	44.98	46.03	46.41	45.21	41.72	42.8	43	43.03	42.09	41.29	42.3	42.98	
25	3	4	42.85	44.96	46.06	46.37	45.15	41.7	42.69	42.85	42.9	41.94	41.14	42.13	42.89	
25	3	5	42.78	44.93	46.09	46.35	45.12	41.66	42.61	42.78	42.8	41.86	40.99	42.04	42.77	
25	4	3	14.92	15.92	16.49	16.58	15.96	14.37	14.87	14.91	14.92	14.57	14.23	14.61	14.92	
25	4	4	14.86	15.91	16.45	16.56	15.94	14.35	14.82	14.86	14.86	14.5	14.14	14.54	14.86	
25	4	5	14.82	15.87	16.47	16.55	15.92	14.35	14.79	14.83	14.82	14.46	14.1	14.5	14.82	
25	5	2	8.72	9.41	9.82	9.86	9.42	8.32	8.71	8.73	8.72	8.51	8.31	8.53	8.72	
25	5	3	8.64	9.36	9.81	9.84	9.38	8.3	8.63	8.64	8.65	8.44	8.22	8.45	8.65	
25	5	4	8.61	9.34	9.79	9.83	9.35	8.3	8.59	8.61	8.61	8.4	8.18	8.41	8.61	
25	5	5	8.58	9.33	9.78	9.81	9.34	8.29	8.57	8.58	8.58	8.37	8.15	8.38	8.58	

Table A.4: Simulation data example for average confidence interval width of σ_{op}^2

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	4.6	4.34	4.14	4.05	4.18	4.06	4.66	4.59	4.57	4.79	4.76	4.67	4.6	
5	3	3	4.7	4.21	3.96	3.83	4.01	3.87	4.85	4.69	4.63	5.03	5.09	4.86	4.7	
5	3	4	4.57	4.02	3.78	3.68	3.83	3.71	4.78	4.58	4.5	5.06	5.16	4.81	4.58	
5	3	5	4.43	3.87	3.65	3.58	3.71	3.61	4.65	4.43	4.34	4.99	5.13	4.7	4.43	
5	4	2	3.36	2.99	2.77	2.7	2.87	2.77	3.43	3.36	3.33	3.55	3.6	3.47	3.37	
5	4	3	3.29	2.75	2.56	2.51	2.66	2.58	3.41	3.29	3.24	3.68	3.8	3.51	3.28	
5	4	4	3.11	2.61	2.47	2.43	2.54	2.48	3.25	3.1	3.06	3.59	3.83	3.39	3.1	
5	5	2	2.79	2.33	2.13	2.09	2.26	2.18	2.86	2.79	2.77	3.02	3.1	2.93	2.79	
5	5	3	2.62	2.11	1.99	1.96	2.07	2.02	2.71	2.62	2.59	3.02	3.25	2.88	2.61	
5	5	4	2.42	2.01	1.92	1.91	1.98	1.95	2.51	2.42	2.4	2.87	3.24	2.72	2.42	
5	5	5	2.28	1.96	1.89	1.88	1.93	1.91	2.36	2.28	2.26	2.71	3.18	2.56	2.28	
10	3	2	2.62	2.13	1.93	1.88	2.02	1.99	2.75	2.62	2.58	2.92	2.97	2.8	2.63	
10	3	3	2.4	1.92	1.81	1.78	1.86	1.84	2.56	2.41	2.36	2.91	3.09	2.69	2.41	
10	3	4	2.21	1.84	1.76	1.74	1.79	1.78	2.34	2.21	2.17	2.74	3.06	2.5	2.21	
10	3	5	2.08	1.79	1.73	1.71	1.76	1.75	2.18	2.07	2.05	2.56	2.99	2.35	2.08	
10	4	2	2.05	1.54	1.42	1.39	1.5	1.48	2.15	2.05	2.02	2.4	2.54	2.28	2.05	
10	4	4	1.62	1.36	1.31	1.3	1.35	1.34	1.67	1.61	1.61	2	2.54	1.88	1.62	
10	4	5	1.53	1.33	1.3	1.29	1.32	1.32	1.57	1.53	1.52	1.84	2.44	1.73	1.53	
10	5	2	1.72	1.26	1.17	1.15	1.24	1.23	1.79	1.72	1.71	2.1	2.32	2	1.72	
10	5	3	1.45	1.17	1.11	1.11	1.16	1.15	1.49	1.45	1.44	1.84	2.34	1.74	1.45	
10	5	4	1.33	1.13	1.09	1.09	1.12	1.12	1.35	1.33	1.32	1.61	2.25	1.54	1.33	
10	5	5	1.26	1.11	1.08	1.07	1.1	1.1	1.28	1.26	1.26	1.47	2.12	1.43	1.26	
15	3	2	2.01	1.5	1.38	1.36	1.46	1.44	2.14	2.01	1.98	2.41	2.54	2.26	2.01	
15	3	3	1.73	1.38	1.31	1.3	1.36	1.35	1.83	1.73	1.71	2.24	2.58	2.03	1.73	
15	3	4	1.58	1.33	1.28	1.27	1.32	1.31	1.64	1.57	1.56	1.99	2.5	1.83	1.57	
15	3	5	1.49	1.3	1.27	1.26	1.29	1.29	1.54	1.49	1.48	1.81	2.4	1.69	1.49	
15	4	2	1.56	1.14	1.06	1.05	1.12	1.12	1.64	1.57	1.55	1.98	2.24	1.87	1.56	
15	4	4	1.2	1.03	0.99	0.99	1.02	1.02	1.23	1.2	1.2	1.46	2.11	1.4	1.2	
15	4	5	1.15	1.01	0.98	0.98	1	1	1.16	1.14	1.14	1.34	1.97	1.29	1.15	
15	5	2	1.31	0.95	0.89	0.88	0.94	0.94	1.35	1.31	1.31	1.71	2.09	1.63	1.31	
15	5	3	1.1	0.89	0.85	0.85	0.89	0.89	1.11	1.1	1.09	1.37	2.03	1.32	1.1	
15	5	4	1.01	0.87	0.84	0.84	0.86	0.86	1.02	1.01	1.01	1.2	1.87	1.17	1.01	
15	5	5	0.96	0.85	0.83	0.83	0.85	0.85	0.97	0.96	0.96	1.1	1.71	1.08	0.96	
20	3	3	1.4	1.13	1.08	1.07	1.12	1.11	1.46	1.4	1.39	1.83	2.31	1.68	1.4	
20	3	4	1.28	1.09	1.06	1.05	1.08	1.08	1.32	1.28	1.28	1.59	2.2	1.49	1.28	
20	3	5	1.22	1.07	1.04	1.04	1.07	1.06	1.25	1.22	1.22	1.45	2.07	1.38	1.22	
20	4	2	1.29	0.94	0.88	0.87	0.93	0.93	1.34	1.3	1.29	1.72	2.09	1.62	1.3	
20	4	3	1.08	0.88	0.85	0.84	0.88	0.88	1.11	1.08	1.08	1.37	2.02	1.31	1.08	
20	4	4	1	0.86	0.83	0.83	0.85	0.85	1.01	1	1	1.19	1.86	1.16	1	
20	5	2	1.09	0.8	0.74	0.74	0.79	0.79	1.11	1.09	1.09	1.46	1.96	1.4	1.09	
20	5	3	0.92	0.75	0.72	0.72	0.75	0.75	0.93	0.92	0.92	1.13	1.84	1.1	0.92	
20	5	4	0.85	0.73	0.7	0.7	0.73	0.73	0.85	0.85	0.85	0.99	1.63	0.98	0.85	
20	5	5	0.81	0.72	0.7	0.7	0.71	0.71	0.81	0.81	0.81	0.92	1.45	0.91	0.81	
25	3	2	1.43	1.05	0.98	0.97	1.03	1.03	1.52	1.44	1.43	1.91	2.19	1.76	1.43	
25	3	3	1.2	0.98	0.94	0.93	0.97	0.97	1.25	1.2	1.2	1.56	2.14	1.45	1.2	
25	3	4	1.11	0.95	0.92	0.91	0.94	0.94	1.13	1.11	1.11	1.35	2	1.28	1.11	
25	3	5	1.06	0.93	0.91	0.9	0.93	0.93	1.07	1.06	1.05	1.23	1.85	1.19	1.06	
25	4	3	0.94	0.77	0.74	0.74	0.77	0.77	0.96	0.94	0.94	1.17	1.87	1.14	0.95	
25	4	4	0.87	0.75	0.73	0.72	0.75	0.75	0.88	0.87	0.87	1.03	1.67	1.01	0.87	
25	4	5	0.83	0.74	0.72	0.72	0.74	0.74	0.84	0.83	0.83	0.95	1.49	0.94	0.83	
25	5	2	0.95	0.7	0.65	0.65	0.7	0.7	0.96	0.95	0.95	1.27	1.86	1.23	0.95	
25	5	3	0.8	0.66	0.63	0.63	0.66	0.66	0.81	0.8	0.8	0.98	1.69	0.96	0.8	
25	5	4	0.74	0.64	0.62	0.62	0.64	0.64	0.75	0.74	0.74	0.86	1.45	0.86	0.74	
25	5	5	0.71	0.63	0.61	0.61	0.63	0.63	0.71	0.71	0.71	0.8	1.27	0.8	0.71	

Table A.5: Simulation data example for average confidence interval width of $\sigma_{Reproducibility}^2$

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	39.85	34.4	31.55	29.97	31.71	32.46	41.42	39.79	39.03	42.96	43.14	41.48	39.9	
5	3	3	36.01	30.89	29.03	28.12	29.1	31.29	38.13	36.05	35.16	41.6	41.65	38.52	36.12	
5	3	4	33.77	29.62	28.39	27.54	28.2	30.87	35.82	33.76	33.06	39.12	40.28	36.25	33.79	
5	3	5	32.51	28.96	27.97	27.21	27.74	30.64	34.32	32.48	31.9	37.59	39.21	34.73	32.51	
5	4	2	12.7	10.04	9.26	8.93	9.52	10.4	13.4	12.7	12.45	14.5	14.89	13.76	12.7	
5	4	3	11.1	9.28	8.81	8.59	8.95	10.15	11.71	11.08	10.89	13.22	14.12	12.27	11.09	
5	4	4	10.46	9.08	8.69	8.48	8.78	10.09	10.94	10.43	10.31	12.31	13.54	11.47	10.46	
5	5	2	6.88	5.29	4.91	4.81	5.11	5.84	7.24	6.86	6.76	8.15	8.63	7.7	6.87	
5	5	3	6.01	5.02	4.77	4.68	4.88	5.74	6.25	6.01	5.95	7.19	8.08	6.74	6	
5	5	4	5.73	4.92	4.69	4.61	4.8	5.71	5.91	5.74	5.69	6.68	7.71	6.33	5.74	
5	5	5	5.61	4.87	4.66	4.58	4.76	5.68	5.78	5.6	5.57	6.41	7.44	6.13	5.61	
10	3	2	29.65	21.43	19.34	18.47	20.02	25.95	32.75	29.74	28.68	38.39	40.27	34.47	29.76	
10	3	3	25.8	20.51	18.86	18.06	19.29	25.61	27.79	25.86	25.26	33.48	37.83	29.75	25.83	
10	3	4	24.82	20.08	18.5	17.8	19.01	25.43	26.16	24.73	24.35	30.51	35.97	27.85	24.72	
10	3	5	24.29	19.82	18.44	17.66	18.83	25.36	25.53	24.31	23.9	29.17	34.51	27.11	24.27	
10	4	2	9	6.5	5.87	5.66	6.24	8.71	9.6	8.99	8.85	12.01	13.71	10.87	9	
10	4	4	8	6.18	5.63	5.48	5.99	8.58	8.3	8	7.93	9.63	12.03	9.13	8.01	
10	4	5	7.9	6.12	5.59	5.44	5.94	8.56	8.15	7.89	7.83	9.39	11.6	8.97	7.9	
10	5	2	4.92	3.54	3.17	3.09	3.45	5.05	5.14	4.92	4.89	6.43	7.87	5.98	4.92	
10	5	3	4.63	3.44	3.1	3.02	3.37	5	4.77	4.63	4.6	5.62	7.25	5.39	4.63	
10	5	4	4.53	3.38	3.06	2.99	3.32	4.98	4.65	4.53	4.52	5.41	6.9	5.23	4.53	
10	5	5	4.48	3.36	3.04	2.98	3.28	4.98	4.59	4.48	4.47	5.32	6.69	5.16	4.48	
15	3	2	24.17	17.19	15.2	14.44	16.29	23.83	26.52	24.17	23.7	33.86	38.44	29.65	24.28	
15	3	3	22.24	16.69	14.79	14.17	15.82	23.66	23.58	22.22	21.98	28.4	35.47	25.93	22.26	
15	3	4	21.7	16.44	14.65	14.03	15.66	23.54	22.8	21.72	21.51	26.69	33.5	24.99	21.68	
15	3	5	21.41	16.33	14.54	13.95	15.55	23.47	22.41	21.46	21.29	26.02	32.34	24.51	21.48	
15	4	2	7.68	5.33	4.63	4.47	5.16	8.18	8.07	7.68	7.62	10.25	13.02	9.43	7.69	
15	4	4	7.21	5.13	4.47	4.36	4.99	8.12	7.45	7.22	7.18	8.76	11.34	8.45	7.21	
15	4	5	7.14	5.09	4.46	4.32	4.97	8.1	7.33	7.13	7.13	8.65	11.02	8.37	7.14	
15	5	2	4.37	2.95	2.53	2.47	2.89	4.81	4.49	4.36	4.35	5.58	7.46	5.31	4.37	
15	5	3	4.22	2.87	2.48	2.42	2.83	4.79	4.3	4.21	4.21	5.15	6.85	5.01	4.21	
15	5	4	4.15	2.83	2.45	2.41	2.79	4.77	4.23	4.15	4.15	5.04	6.57	4.93	4.16	
15	5	5	4.12	2.82	2.44	2.39	2.78	4.76	4.2	4.13	4.12	4.98	6.44	4.89	4.12	
20	3	3	20.64	14.66	12.68	12.14	14.06	22.67	21.67	20.55	20.4	25.86	33.98	24.11	20.56	
20	3	4	20.24	14.47	12.58	12.02	13.93	22.62	21.13	20.26	20.1	24.88	32.05	23.57	20.27	
20	3	5	20.1	14.34	12.51	11.96	13.86	22.55	20.83	20.03	19.89	24.45	30.94	23.31	20.05	
20	4	2	7.12	4.72	3.98	3.88	4.62	7.95	7.38	7.13	7.11	9.25	12.57	8.71	7.14	
20	4	3	6.92	4.61	3.91	3.8	4.53	7.91	7.1	6.94	6.91	8.55	11.47	8.24	6.93	
20	4	4	6.83	4.57	3.88	3.76	4.49	7.89	6.99	6.82	6.79	8.36	10.97	8.14	6.83	
20	5	2	4.13	2.64	2.19	2.16	2.61	4.7	4.22	4.12	4.12	5.18	7.2	5.02	4.13	
20	5	3	4.02	2.58	2.16	2.12	2.56	4.68	4.09	4.02	4.02	4.93	6.63	4.84	4.02	
20	5	4	3.98	2.55	2.14	2.1	2.53	4.67	4.02	3.98	3.98	4.87	6.41	4.78	3.97	
20	5	5	3.95	2.54	2.12	2.09	2.53	4.66	4.01	3.95	3.96	4.82	6.33	4.75	3.95	
25	3	2	20.25	13.74	11.63	11.1	13.19	22.2	21.42	20.31	20.17	27.61	36.08	24.91	20.28	
25	3	3	19.62	13.44	11.46	10.88	12.99	22.11	20.44	19.6	19.5	24.49	32.82	23.21	19.64	
25	3	4	19.36	13.27	11.27	10.83	12.86	21.97	20.02	19.36	19.25	23.87	31.01	22.81	19.29	
25	3	5	19.18	13.16	11.14	10.76	12.8	21.99	19.81	19.2	19.16	23.5	30.18	22.61	19.21	
25	4	3	6.68	4.28	3.54	3.45	4.22	7.77	6.82	6.69	6.67	8.25	11.15	8.04	6.68	
25	4	4	6.62	4.23	3.53	3.42	4.18	7.76	6.73	6.61	6.61	8.13	10.74	7.94	6.61	
25	4	5	6.57	4.22	3.49	3.41	4.15	7.75	6.69	6.57	6.58	8.06	10.58	7.89	6.57	
25	5	2	3.99	2.46	2	1.97	2.44	4.63	4.04	3.98	3.99	4.98	7.01	4.87	3.99	
25	5	3	3.92	2.42	1.96	1.94	2.4	4.62	3.95	3.91	3.9	4.82	6.48	4.75	3.91	
25	5	4	3.88	2.4	1.95	1.92	2.38	4.61	3.91	3.88	3.87	4.74	6.32	4.7	3.88	
25	5	5	3.85	2.38	1.94	1.92	2.37	4.61	3.89	3.85	3.86	4.72	6.27	4.67	3.86	

Table A.6: Simulation data example for average confidence interval width of σ_{GRR}^2

p o n			σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	23.52	24.07	24.36	24.03	23.72	27.85	23.49	23.5	23.46	23.26	22.92	23.26	23.48	
5	3	3	22.09	23.03	23.35	23.35	22.8	27.37	22.22	22.09	22.07	21.63	21.08	21.7	22.05	
5	3	4	21.29	22.49	23.16	22.99	22.4	27.11	21.23	21.32	21.35	20.83	20.24	20.96	21.34	
5	3	5	20.89	22.12	22.96	22.76	22.08	26.92	20.81	20.89	20.97	20.28	19.67	20.43	20.94	
5	4	2	7.48	7.5	7.53	7.45	7.39	9.06	7.55	7.46	7.44	7.55	7.48	7.48	7.47	
5	4	3	7.04	7.21	7.31	7.23	7.13	8.91	7.08	7.03	7.01	7.09	6.96	6.99	7.04	
5	4	4	6.84	7.08	7.22	7.16	7.03	8.86	6.86	6.81	6.82	6.81	6.69	6.77	6.82	
5	5	2	4.11	4.08	4.08	4.05	4.03	5.1	4.15	4.12	4.1	4.21	4.22	4.15	4.11	
5	5	3	3.89	3.93	3.98	3.95	3.89	5.03	3.93	3.9	3.88	3.95	3.95	3.9	3.89	
5	5	4	3.78	3.87	3.92	3.89	3.83	5	3.79	3.78	3.78	3.81	3.81	3.77	3.78	
5	5	5	3.71	3.83	3.9	3.88	3.81	4.98	3.74	3.71	3.71	3.73	3.73	3.7	3.72	
10	3	2	17.79	16.65	16.15	15.72	16.1	22.83	18.21	17.77	17.59	18.89	18.96	18.27	17.74	
10	3	3	16.91	16.17	15.8	15.42	15.62	22.58	17.32	16.93	16.78	17.8	17.9	17.27	16.93	
10	3	4	16.55	15.86	15.5	15.2	15.42	22.44	16.8	16.5	16.43	17.15	17.4	16.84	16.48	
10	3	5	16.28	15.66	15.48	15.09	15.27	22.39	16.53	16.34	16.18	16.84	17.02	16.61	16.28	
10	4	2	5.68	5.13	4.95	4.82	5	7.65	5.78	5.67	5.64	6.12	6.33	5.93	5.68	
10	4	4	5.33	4.9	4.75	4.68	4.83	7.55	5.42	5.34	5.32	5.66	5.92	5.55	5.35	
10	4	5	5.29	4.86	4.72	4.64	4.79	7.53	5.34	5.27	5.26	5.57	5.83	5.49	5.28	
10	5	2	3.17	2.79	2.67	2.63	2.75	4.43	3.23	3.17	3.17	3.43	3.65	3.36	3.18	
10	5	3	3.07	2.73	2.62	2.57	2.69	4.4	3.11	3.07	3.06	3.3	3.53	3.24	3.07	
10	5	4	3.01	2.68	2.58	2.54	2.65	4.37	3.05	3.01	3.01	3.23	3.46	3.18	3.01	
10	5	5	2.98	2.66	2.57	2.53	2.63	4.37	3.02	2.98	2.98	3.18	3.41	3.14	2.98	
15	3	2	15.48	13.62	12.84	12.39	13.21	21.03	15.93	15.49	15.42	16.94	17.49	16.35	15.55	
15	3	3	14.93	13.28	12.52	12.17	12.84	20.9	15.32	14.89	14.87	16.06	16.76	15.69	14.93	
15	3	4	14.65	13.11	12.42	12.06	12.73	20.8	15	14.69	14.6	15.75	16.39	15.42	14.64	
15	3	5	14.5	13.05	12.32	11.99	12.64	20.74	14.77	14.52	14.49	15.53	16.23	15.15	14.54	
15	4	2	5.05	4.23	3.92	3.82	4.15	7.2	5.18	5.04	5.03	5.55	5.95	5.42	5.06	
15	4	4	4.83	4.1	3.79	3.73	4.02	7.15	4.92	4.84	4.82	5.25	5.69	5.18	4.83	
15	4	5	4.79	4.07	3.78	3.7	4	7.14	4.85	4.79	4.79	5.21	5.64	5.14	4.79	
15	5	2	2.88	2.34	2.14	2.11	2.3	4.23	2.92	2.87	2.87	3.18	3.48	3.12	2.88	
15	5	3	2.81	2.28	2.1	2.06	2.26	4.21	2.83	2.8	2.81	3.09	3.4	3.05	2.8	
15	5	4	2.77	2.25	2.08	2.06	2.23	4.2	2.8	2.77	2.76	3.04	3.36	3.01	2.78	
15	5	5	2.75	2.25	2.07	2.04	2.22	4.19	2.78	2.75	2.75	3.01	3.33	2.99	2.75	
20	3	3	13.98	11.75	10.78	10.46	11.42	20.05	14.32	13.9	13.84	15.17	16.17	14.82	13.9	
20	3	4	13.74	11.6	10.71	10.35	11.32	20	14.03	13.74	13.68	14.9	15.92	14.59	13.76	
20	3	5	13.67	11.51	10.65	10.31	11.27	19.94	13.88	13.6	13.53	14.76	15.73	14.47	13.62	
20	4	2	4.75	3.77	3.38	3.32	3.72	7.01	4.84	4.74	4.74	5.28	5.76	5.17	4.76	
20	4	3	4.64	3.69	3.33	3.26	3.65	6.98	4.71	4.66	4.64	5.15	5.65	5.05	4.65	
20	4	4	4.59	3.67	3.3	3.23	3.62	6.96	4.66	4.58	4.57	5.06	5.6	5	4.59	
20	5	2	2.74	2.1	1.86	1.84	2.09	4.14	2.78	2.73	2.73	3.05	3.41	3.01	2.74	
20	5	3	2.69	2.06	1.83	1.81	2.05	4.12	2.72	2.68	2.68	2.98	3.34	2.95	2.68	
20	5	4	2.66	2.04	1.82	1.79	2.03	4.12	2.68	2.66	2.66	2.96	3.32	2.93	2.66	
20	5	5	2.64	2.03	1.81	1.79	2.03	4.11	2.67	2.64	2.65	2.93	3.3	2.91	2.64	
25	3	2	13.6	11.03	9.91	9.57	10.7	19.63	13.98	13.65	13.58	15.2	16.26	14.71	13.61	
25	3	3	13.31	10.81	9.79	9.39	10.56	19.56	13.59	13.3	13.27	14.68	15.85	14.36	13.33	
25	3	4	13.17	10.69	9.63	9.35	10.45	19.43	13.36	13.17	13.1	14.47	15.55	14.16	13.11	
25	3	5	13.05	10.6	9.52	9.29	10.41	19.46	13.27	13.07	13.07	14.29	15.53	14.08	13.09	
25	4	3	4.49	3.43	3.02	2.96	3.4	6.86	4.54	4.5	4.48	4.99	5.57	4.94	4.49	
25	4	4	4.46	3.39	3.01	2.94	3.36	6.86	4.5	4.45	4.45	4.96	5.52	4.89	4.45	
25	4	5	4.43	3.38	2.98	2.92	3.34	6.84	4.48	4.43	4.44	4.93	5.5	4.86	4.43	
25	5	2	2.66	1.96	1.7	1.68	1.95	4.08	2.67	2.65	2.66	2.98	3.36	2.95	2.65	
25	5	3	2.62	1.94	1.67	1.66	1.93	4.07	2.63	2.62	2.61	2.93	3.32	2.91	2.62	
25	5	4	2.6	1.92	1.66	1.64	1.91	4.06	2.61	2.6	2.6	2.89	3.29	2.88	2.6	
25	5	5	2.58	1.91	1.65	1.64	1.9	4.07	2.59	2.58	2.59	2.88	3.29	2.87	2.59	

Table A.7: Simulation data example for average confidence interval width of σ_{Total}^2

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	15.67	15.66	15.69	14.27	12.85	17.56	17.03	15.68	14.64	17.76	15.64	15.62	15.64	
5	3	3	15.12	15.35	15.23	14.06	12.59	17.43	16.54	15.1	14.15	17	14.8	14.95	15.07	
5	3	4	14.78	15.1	15.23	13.94	12.52	17.36	16.05	14.79	13.91	16.57	14.44	14.71	14.83	
5	3	5	14.62	14.91	15.22	13.86	12.36	17.27	15.89	14.62	13.8	16.33	14.17	14.48	14.69	
5	4	2	6.13	6.14	6.15	6.04	6.02	6.99	6.32	6.12	6.05	6.45	6.17	6.13	6.12	
5	4	3	5.91	5.98	6.04	5.93	5.93	6.92	6.06	5.9	5.85	6.16	5.88	5.88	5.91	
5	4	4	5.8	5.92	5.98	5.89	5.89	6.89	5.93	5.78	5.76	5.99	5.74	5.77	5.78	
5	5	2	4.21	4.17	4.18	4.41	4.81	4.9	4.05	4.21	4.4	4.04	4.31	4.24	4.2	
5	5	3	4.06	4.07	4.1	4.35	4.76	4.86	3.87	4.06	4.27	3.83	4.13	4.08	4.06	
5	5	4	3.98	4.02	4.06	4.32	4.72	4.84	3.77	3.98	4.21	3.72	4.03	3.99	3.98	
5	5	5	3.94	3.99	4.03	4.3	4.71	4.83	3.73	3.93	4.17	3.65	3.98	3.94	3.93	
10	3	2	12.71	11.89	11.41	10.12	8.92	15.11	14.16	12.7	11.57	15.28	13.36	12.98	12.65	
10	3	3	12.35	11.67	11.27	10.01	8.73	14.96	13.86	12.36	11.25	14.81	12.8	12.54	12.36	
10	3	4	12.18	11.51	11.07	9.9	8.65	14.91	13.57	12.14	11.07	14.48	12.55	12.36	12.14	
10	3	5	12.04	11.38	11.11	9.87	8.55	14.93	13.45	12.09	10.96	14.33	12.32	12.26	12.06	
10	4	2	4.36	4.02	3.89	3.53	3.26	5.32	4.81	4.36	4.03	5.23	4.67	4.49	4.37	
10	4	4	4.17	3.89	3.77	3.45	3.18	5.27	4.61	4.18	3.86	4.98	4.44	4.29	4.18	
10	4	5	4.15	3.87	3.76	3.43	3.16	5.26	4.57	4.14	3.83	4.93	4.39	4.25	4.14	
10	5	2	2.62	2.38	2.29	2.19	2.21	3.27	2.82	2.62	2.5	3.05	2.87	2.72	2.62	
10	5	3	2.55	2.34	2.26	2.16	2.19	3.25	2.75	2.56	2.44	2.97	2.79	2.65	2.55	
10	5	4	2.51	2.3	2.24	2.14	2.17	3.24	2.71	2.52	2.41	2.92	2.75	2.61	2.52	
10	5	5	2.5	2.29	2.23	2.14	2.16	3.24	2.69	2.49	2.4	2.89	2.72	2.58	2.49	
15	3	2	11.51	10.22	9.58	8.41	7.41	14.09	12.97	11.53	10.44	14.21	12.48	11.99	11.58	
15	3	3	11.26	10.06	9.41	8.3	7.21	14.04	12.73	11.22	10.13	13.82	12.04	11.67	11.24	
15	3	4	11.1	9.99	9.36	8.24	7.17	13.97	12.56	11.14	9.97	13.71	11.82	11.54	11.09	
15	3	5	11.03	9.98	9.32	8.22	7.12	13.93	12.42	11.05	9.92	13.6	11.73	11.34	11.06	
15	4	2	3.92	3.4	3.17	2.83	2.59	4.93	4.43	3.92	3.55	4.86	4.33	4.11	3.93	
15	4	4	3.79	3.33	3.09	2.78	2.53	4.9	4.28	3.8	3.42	4.71	4.18	3.97	3.79	
15	4	5	3.76	3.3	3.09	2.76	2.52	4.9	4.23	3.76	3.41	4.69	4.14	3.94	3.76	
15	5	2	2.32	1.98	1.84	1.71	1.7	2.98	2.57	2.31	2.15	2.86	2.62	2.45	2.33	
15	5	3	2.27	1.94	1.81	1.68	1.68	2.98	2.52	2.27	2.11	2.81	2.57	2.4	2.27	
15	5	4	2.25	1.92	1.8	1.67	1.66	2.97	2.49	2.25	2.08	2.77	2.54	2.38	2.26	
15	5	5	2.23	1.92	1.79	1.66	1.66	2.97	2.49	2.24	2.08	2.76	2.52	2.36	2.24	
20	3	3	10.7	9.17	8.38	7.33	6.39	13.48	12.16	10.64	9.47	13.32	11.6	11.11	10.63	
20	3	4	10.55	9.09	8.36	7.27	6.34	13.43	12	10.55	9.37	13.19	11.45	10.95	10.57	
20	3	5	10.53	9.04	8.33	7.24	6.31	13.4	11.91	10.47	9.27	13.14	11.33	10.88	10.48	
20	4	2	3.7	3.08	2.79	2.47	2.27	4.76	4.2	3.69	3.32	4.7	4.16	3.92	3.7	
20	4	3	3.63	3.03	2.76	2.43	2.23	4.74	4.13	3.64	3.25	4.64	4.1	3.84	3.64	
20	4	4	3.6	3.02	2.74	2.41	2.22	4.73	4.1	3.59	3.2	4.59	4.07	3.8	3.59	
20	5	2	2.19	1.79	1.61	1.47	1.46	2.87	2.47	2.18	1.99	2.77	2.51	2.34	2.19	
20	5	3	2.15	1.76	1.59	1.45	1.45	2.86	2.43	2.15	1.96	2.73	2.47	2.29	2.14	
20	5	4	2.13	1.74	1.58	1.44	1.44	2.85	2.4	2.13	1.95	2.72	2.46	2.27	2.13	
20	5	5	2.12	1.74	1.57	1.44	1.44	2.85	2.4	2.12	1.94	2.7	2.45	2.26	2.12	
25	3	2	10.42	8.72	7.85	6.79	5.94	13.15	11.9	10.47	9.25	13.28	11.6	10.99	10.43	
25	3	3	10.26	8.6	7.79	6.68	5.88	13.11	11.7	10.24	9.07	13.06	11.35	10.78	10.28	
25	3	4	10.17	8.54	7.69	6.65	5.82	13.01	11.58	10.18	8.95	12.97	11.16	10.64	10.13	
25	3	5	10.09	8.48	7.6	6.62	5.79	13.04	11.55	10.11	8.94	12.85	11.18	10.59	10.12	
25	4	3	3.51	2.85	2.54	2.21	2.05	4.62	4.02	3.51	3.11	4.54	4.02	3.74	3.51	
25	4	4	3.48	2.82	2.54	2.2	2.03	4.63	4	3.48	3.09	4.53	3.99	3.7	3.48	
25	4	5	3.46	2.82	2.51	2.19	2.02	4.62	3.99	3.47	3.09	4.52	3.98	3.68	3.47	
25	5	2	2.11	1.67	1.48	1.33	1.33	2.8	2.39	2.1	1.9	2.72	2.46	2.26	2.1	
25	5	3	2.08	1.66	1.46	1.31	1.31	2.79	2.37	2.08	1.87	2.7	2.43	2.23	2.08	
25	5	4	2.06	1.64	1.45	1.31	1.31	2.78	2.35	2.06	1.87	2.67	2.41	2.21	2.06	
25	5	5	2.05	1.63	1.45	1.3	1.3	2.79	2.34	2.05	1.86	2.67	2.41	2.2	2.06	

Table A.8: Simulation data example for average confidence interval width of ρ_p

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	10.89	11.1	11.24	11.01	10.63	11.11	11.24	10.9	10.7	11.67	10.93	10.86	10.89	
5	3	3	10.79	11.13	11.24	10.99	10.59	11.08	11.01	10.79	10.62	11.19	10.52	10.63	10.78	
5	3	4	10.78	11.12	11.23	10.97	10.58	11.07	10.98	10.78	10.61	10.98	10.32	10.59	10.79	
5	3	5	10.77	11.11	11.23	10.96	10.55	11.05	11	10.78	10.6	10.94	10.21	10.58	10.79	
5	4	2	10.6	11.03	11.17	10.75	10.22	10.64	10.97	10.6	10.35	11.16	10.35	10.46	10.61	
5	4	3	10.61	11.02	11.17	10.72	10.17	10.59	11.01	10.6	10.32	10.95	10.04	10.36	10.6	
5	4	4	10.6	11.01	11.13	10.71	10.15	10.57	11.03	10.59	10.29	10.96	9.89	10.34	10.59	
5	5	2	10.43	10.96	11.1	10.54	9.92	10.28	10.95	10.44	10.11	10.98	10	10.21	10.44	
5	5	3	10.42	10.92	11.09	10.52	9.86	10.23	11.01	10.43	10.07	10.94	9.73	10.15	10.42	
5	5	4	10.39	10.9	11.05	10.48	9.85	10.21	11.01	10.38	10.02	10.99	9.61	10.11	10.38	
5	5	5	10.35	10.89	11.06	10.48	9.83	10.2	11	10.35	9.98	11.03	9.55	10.07	10.35	
10	3	2	5.1	5.3	5.33	5.04	4.69	5.22	5.34	5.1	4.9	5.38	4.85	4.98	5.09	
10	3	3	5.06	5.26	5.3	5.01	4.64	5.17	5.38	5.06	4.84	5.34	4.66	4.92	5.06	
10	3	4	5.02	5.23	5.27	5	4.62	5.16	5.35	5.02	4.79	5.36	4.58	4.87	5.02	
10	3	5	4.99	5.21	5.27	4.99	4.6	5.15	5.34	4.99	4.76	5.36	4.53	4.84	4.99	
10	4	2	4.8	5.06	5.14	4.74	4.28	4.72	5.21	4.8	4.52	5.22	4.44	4.64	4.8	
10	4	4	4.68	5.01	5.1	4.7	4.22	4.67	5.17	4.68	4.39	5.25	4.22	4.5	4.68	
10	4	5	4.65	5	5.1	4.69	4.21	4.66	5.15	4.65	4.36	5.22	4.17	4.46	4.64	
10	5	2	4.54	4.89	5	4.49	4	4.39	5.09	4.55	4.24	5.13	4.17	4.38	4.54	
10	5	3	4.46	4.85	4.99	4.47	3.96	4.36	5.05	4.46	4.16	5.14	4.04	4.29	4.46	
10	5	4	4.41	4.83	4.97	4.45	3.95	4.34	5.01	4.41	4.11	5.11	3.98	4.23	4.41	
10	5	5	4.38	4.82	4.97	4.44	3.94	4.34	4.99	4.38	4.09	5.08	3.94	4.18	4.38	
15	3	2	3.95	4.1	4.12	3.81	3.45	3.98	4.28	3.95	3.71	4.28	3.67	3.84	3.95	
15	3	3	3.88	4.07	4.1	3.79	3.4	3.95	4.27	3.88	3.63	4.32	3.53	3.76	3.88	
15	3	4	3.84	4.05	4.09	3.78	3.39	3.93	4.24	3.84	3.58	4.32	3.46	3.7	3.84	
15	3	5	3.81	4.05	4.08	3.77	3.37	3.92	4.21	3.81	3.56	4.3	3.42	3.66	3.81	
15	4	2	3.62	3.88	3.96	3.51	3.08	3.55	4.12	3.61	3.33	4.18	3.31	3.49	3.62	
15	4	4	3.49	3.84	3.93	3.48	3.03	3.51	4.04	3.49	3.22	4.14	3.15	3.33	3.49	
15	4	5	3.46	3.83	3.93	3.47	3.02	3.5	4.01	3.46	3.2	4.11	3.11	3.3	3.46	
15	5	2	3.35	3.7	3.83	3.27	2.84	3.27	3.96	3.35	3.07	4.07	3.07	3.22	3.35	
15	5	3	3.27	3.66	3.81	3.25	2.82	3.25	3.89	3.27	3.01	4.02	2.99	3.13	3.27	
15	5	4	3.22	3.64	3.8	3.23	2.8	3.24	3.86	3.22	2.98	3.97	2.95	3.08	3.23	
15	5	5	3.2	3.63	3.79	3.22	2.79	3.23	3.84	3.2	2.96	3.94	2.92	3.06	3.2	
20	3	3	3.33	3.54	3.57	3.23	2.84	3.39	3.77	3.33	3.07	3.86	3.02	3.21	3.33	
20	3	4	3.29	3.52	3.56	3.22	2.83	3.37	3.73	3.29	3.03	3.83	2.97	3.15	3.29	
20	3	5	3.26	3.52	3.56	3.21	2.82	3.36	3.72	3.26	3.01	3.81	2.93	3.11	3.26	
20	4	2	3.06	3.34	3.43	2.94	2.54	3.03	3.62	3.06	2.79	3.73	2.81	2.95	3.06	
20	4	3	2.97	3.31	3.41	2.92	2.52	3.01	3.56	2.98	2.73	3.68	2.73	2.86	2.98	
20	4	4	2.94	3.3	3.41	2.91	2.51	3	3.52	2.94	2.7	3.64	2.69	2.81	2.93	
20	5	2	2.8	3.15	3.29	2.7	2.33	2.79	3.44	2.8	2.56	3.58	2.61	2.7	2.8	
20	5	3	2.72	3.11	3.28	2.67	2.31	2.77	3.37	2.73	2.51	3.51	2.54	2.63	2.72	
20	5	4	2.69	3.09	3.27	2.66	2.3	2.77	3.33	2.69	2.49	3.46	2.51	2.59	2.69	
20	5	5	2.66	3.08	3.26	2.65	2.3	2.76	3.31	2.67	2.47	3.43	2.49	2.57	2.66	
25	3	2	3.07	3.25	3.28	2.91	2.54	3.08	3.52	3.07	2.81	3.6	2.82	2.97	3.07	
25	3	3	2.99	3.22	3.27	2.89	2.51	3.06	3.47	2.99	2.74	3.59	2.73	2.88	2.99	
25	3	4	2.95	3.21	3.26	2.88	2.5	3.04	3.44	2.95	2.7	3.56	2.68	2.82	2.95	
25	3	5	2.92	3.2	3.25	2.87	2.49	3.04	3.42	2.92	2.69	3.53	2.66	2.79	2.92	
25	4	3	2.65	2.99	3.11	2.57	2.22	2.72	3.25	2.65	2.43	3.39	2.47	2.55	2.64	
25	4	4	2.61	2.97	3.1	2.56	2.21	2.71	3.22	2.61	2.4	3.34	2.43	2.51	2.61	
25	4	5	2.58	2.96	3.09	2.55	2.2	2.71	3.2	2.59	2.39	3.32	2.41	2.48	2.58	
25	5	2	2.47	2.81	2.98	2.35	2.05	2.52	3.12	2.47	2.27	3.29	2.35	2.4	2.47	
25	5	3	2.41	2.78	2.96	2.32	2.03	2.51	3.05	2.41	2.23	3.21	2.3	2.34	2.41	
25	5	4	2.38	2.76	2.95	2.32	2.02	2.5	3.01	2.38	2.21	3.15	2.27	2.31	2.38	
25	5	5	2.36	2.75	2.94	2.31	2.02	2.5	2.98	2.36	2.2	3.12	2.25	2.29	2.36	

Table A.9: Simulation data example for average confidence interval width of ρ_{Rep}

			σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
p	o	n	σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	2.11	2.49	2.72	2.79	2.63	3.11	2.02	2.11	2.15	1.93	1.91	2	2.1	
5	3	3	1.98	2.23	2.34	2.38	2.29	2.68	1.92	1.98	2	1.8	1.75	1.87	1.97	
5	3	4	1.9	2.11	2.21	2.24	2.15	2.53	1.85	1.9	1.92	1.75	1.69	1.82	1.9	
5	3	5	1.86	2.04	2.15	2.16	2.08	2.45	1.82	1.86	1.87	1.72	1.66	1.78	1.86	
5	4	2	1.87	2.19	2.3	2.32	2.23	2.57	1.8	1.87	1.89	1.68	1.63	1.75	1.87	
5	4	3	1.72	1.89	1.96	1.96	1.9	2.21	1.69	1.72	1.73	1.59	1.52	1.64	1.72	
5	4	4	1.64	1.78	1.84	1.85	1.79	2.1	1.62	1.63	1.64	1.54	1.47	1.58	1.64	
5	5	2	1.68	1.91	1.97	1.98	1.92	2.21	1.64	1.68	1.69	1.51	1.44	1.57	1.68	
5	5	3	1.51	1.63	1.68	1.68	1.63	1.92	1.5	1.51	1.51	1.42	1.35	1.45	1.51	
5	5	4	1.43	1.53	1.58	1.59	1.54	1.82	1.42	1.43	1.43	1.37	1.31	1.38	1.42	
5	5	5	1.38	1.48	1.54	1.54	1.49	1.78	1.37	1.38	1.38	1.33	1.28	1.34	1.38	
10	3	2	1.87	2.03	2.07	2.08	2.04	2.42	1.82	1.87	1.89	1.7	1.66	1.77	1.87	
10	3	3	1.72	1.8	1.83	1.83	1.8	2.17	1.71	1.72	1.72	1.65	1.59	1.68	1.72	
10	3	4	1.65	1.71	1.74	1.75	1.72	2.08	1.64	1.65	1.65	1.61	1.56	1.62	1.65	
10	3	5	1.61	1.67	1.71	1.71	1.68	2.05	1.6	1.61	1.61	1.58	1.53	1.59	1.61	
10	4	2	1.61	1.67	1.69	1.69	1.67	2.02	1.59	1.61	1.61	1.51	1.44	1.55	1.61	
10	4	4	1.38	1.41	1.43	1.43	1.41	1.77	1.38	1.38	1.38	1.38	1.36	1.38	1.38	
10	4	5	1.35	1.38	1.4	1.4	1.38	1.74	1.35	1.35	1.35	1.35	1.34	1.35	1.35	
10	5	2	1.4	1.42	1.43	1.42	1.41	1.77	1.4	1.4	1.4	1.36	1.29	1.37	1.4	
10	5	3	1.25	1.25	1.26	1.26	1.25	1.61	1.26	1.25	1.25	1.26	1.24	1.26	1.25	
10	5	4	1.2	1.2	1.21	1.2	1.19	1.56	1.2	1.2	1.2	1.21	1.21	1.2	1.2	
10	5	5	1.17	1.17	1.18	1.18	1.17	1.54	1.17	1.17	1.17	1.18	1.2	1.18	1.17	
15	3	2	1.73	1.77	1.78	1.78	1.77	2.15	1.72	1.73	1.74	1.64	1.58	1.68	1.74	
15	3	3	1.58	1.6	1.61	1.61	1.6	1.98	1.59	1.58	1.58	1.57	1.53	1.58	1.58	
15	3	4	1.52	1.54	1.55	1.55	1.54	1.92	1.53	1.53	1.52	1.53	1.5	1.53	1.52	
15	3	5	1.49	1.51	1.52	1.52	1.51	1.89	1.49	1.49	1.5	1.5	1.49	1.49	1.5	
15	4	2	1.46	1.44	1.44	1.44	1.44	1.82	1.46	1.46	1.46	1.44	1.38	1.45	1.46	
15	4	4	1.28	1.26	1.25	1.25	1.26	1.65	1.28	1.28	1.28	1.3	1.31	1.29	1.28	
15	4	5	1.25	1.23	1.23	1.23	1.23	1.63	1.26	1.25	1.25	1.28	1.3	1.27	1.25	
15	5	2	1.26	1.22	1.2	1.2	1.21	1.61	1.27	1.26	1.26	1.28	1.24	1.28	1.27	
15	5	3	1.15	1.1	1.08	1.08	1.1	1.51	1.15	1.15	1.15	1.18	1.2	1.18	1.15	
15	5	4	1.11	1.06	1.04	1.04	1.05	1.47	1.11	1.11	1.11	1.14	1.18	1.14	1.11	
15	5	5	1.09	1.04	1.02	1.02	1.03	1.45	1.09	1.09	1.09	1.12	1.17	1.12	1.09	
20	3	3	1.5	1.49	1.49	1.48	1.49	1.88	1.51	1.5	1.5	1.51	1.49	1.51	1.5	
20	3	4	1.46	1.44	1.44	1.44	1.44	1.83	1.46	1.46	1.45	1.47	1.47	1.46	1.46	
20	3	5	1.43	1.42	1.42	1.41	1.42	1.81	1.43	1.43	1.43	1.44	1.45	1.44	1.43	
20	4	2	1.37	1.32	1.3	1.3	1.31	1.72	1.37	1.37	1.37	1.38	1.34	1.38	1.37	
20	4	3	1.26	1.21	1.19	1.18	1.2	1.62	1.26	1.26	1.26	1.29	1.31	1.29	1.26	
20	4	4	1.22	1.17	1.15	1.15	1.17	1.59	1.23	1.22	1.22	1.25	1.29	1.25	1.22	
20	5	2	1.19	1.11	1.08	1.07	1.11	1.53	1.19	1.19	1.19	1.22	1.21	1.22	1.19	
20	5	3	1.09	1.01	0.98	0.98	1.01	1.45	1.1	1.09	1.09	1.14	1.18	1.13	1.09	
20	5	4	1.06	0.97	0.95	0.94	0.97	1.42	1.06	1.06	1.06	1.11	1.16	1.1	1.06	
20	5	5	1.04	0.96	0.93	0.93	0.96	1.41	1.05	1.05	1.05	1.09	1.15	1.09	1.05	
25	3	2	1.56	1.53	1.52	1.51	1.52	1.92	1.57	1.56	1.56	1.56	1.51	1.57	1.56	
25	3	3	1.45	1.42	1.41	1.4	1.41	1.81	1.45	1.45	1.45	1.47	1.47	1.47	1.45	
25	3	4	1.41	1.38	1.37	1.36	1.37	1.77	1.41	1.41	1.41	1.43	1.45	1.43	1.41	
25	3	5	1.39	1.36	1.34	1.34	1.35	1.76	1.39	1.39	1.39	1.41	1.44	1.41	1.39	
25	4	3	1.22	1.14	1.11	1.11	1.14	1.58	1.22	1.22	1.22	1.26	1.29	1.25	1.22	
25	4	4	1.19	1.11	1.08	1.08	1.1	1.55	1.19	1.18	1.18	1.23	1.27	1.22	1.18	
25	4	5	1.17	1.09	1.06	1.06	1.09	1.53	1.17	1.17	1.17	1.21	1.26	1.2	1.17	
25	5	2	1.14	1.04	1	0.99	1.03	1.49	1.14	1.14	1.14	1.18	1.2	1.18	1.14	
25	5	3	1.06	0.96	0.91	0.91	0.95	1.42	1.06	1.06	1.06	1.11	1.17	1.11	1.06	
25	5	4	1.03	0.93	0.88	0.88	0.92	1.39	1.04	1.03	1.03	1.08	1.15	1.08	1.03	
25	5	5	1.02	0.91	0.87	0.87	0.91	1.38	1.02	1.02	1.02	1.07	1.14	1.07	1.02	

Table A.10: Simulation data example for average confidence interval width of $\rho_{Reproducibility}$

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	1.4	0.95	0.71	0.61	0.78	0.47	1.53	1.4	1.35	1.7	1.75	1.56	1.41	
5	3	3	1.25	0.74	0.56	0.47	0.61	0.38	1.4	1.25	1.18	1.89	1.86	1.49	1.26	
5	3	4	1.14	0.67	0.5	0.44	0.55	0.35	1.33	1.14	1.08	1.67	1.87	1.39	1.14	
5	3	5	1.08	0.63	0.47	0.42	0.53	0.34	1.25	1.07	1.01	1.62	1.87	1.33	1.07	
5	4	2	1.23	0.68	0.49	0.43	0.59	0.36	1.35	1.23	1.19	1.57	1.68	1.44	1.23	
5	4	3	0.99	0.53	0.39	0.35	0.47	0.3	1.11	0.99	0.95	1.44	1.68	1.26	0.99	
5	4	4	0.88	0.49	0.36	0.33	0.44	0.29	0.99	0.88	0.85	1.32	1.64	1.13	0.89	
5	5	2	1.06	0.53	0.37	0.35	0.48	0.31	1.16	1.06	1.03	1.42	1.57	1.31	1.06	
5	5	3	0.82	0.43	0.31	0.29	0.4	0.26	0.89	0.82	0.8	1.2	1.5	1.07	0.82	
5	5	4	0.74	0.4	0.29	0.27	0.37	0.25	0.8	0.74	0.73	1.07	1.44	0.96	0.74	
5	5	5	0.71	0.39	0.28	0.26	0.36	0.24	0.76	0.71	0.7	1	1.38	0.9	0.71	
10	3	2	1.26	0.59	0.41	0.36	0.5	0.34	1.49	1.27	1.2	1.92	2.1	1.64	1.28	
10	3	3	0.97	0.5	0.36	0.31	0.43	0.3	1.12	0.97	0.93	1.64	2.1	1.33	0.97	
10	3	4	0.89	0.47	0.34	0.3	0.41	0.29	1	0.89	0.86	1.43	2.02	1.18	0.89	
10	3	5	0.86	0.46	0.33	0.29	0.4	0.28	0.96	0.85	0.83	1.32	1.94	1.12	0.86	
10	4	2	0.96	0.44	0.31	0.28	0.41	0.28	1.08	0.96	0.93	1.54	1.87	1.33	0.95	
10	4	4	0.73	0.37	0.26	0.24	0.34	0.25	0.78	0.72	0.71	1.05	1.61	0.96	0.72	
10	4	5	0.7	0.36	0.25	0.23	0.33	0.24	0.75	0.7	0.69	1.01	1.52	0.92	0.7	
10	5	2	0.78	0.37	0.25	0.24	0.35	0.25	0.84	0.78	0.77	1.22	1.62	1.09	0.78	
10	5	3	0.66	0.32	0.22	0.21	0.3	0.22	0.69	0.65	0.65	0.94	1.44	0.87	0.66	
10	5	4	0.62	0.31	0.21	0.2	0.29	0.22	0.65	0.62	0.62	0.87	1.33	0.82	0.62	
10	5	5	0.61	0.3	0.2	0.19	0.28	0.21	0.63	0.61	0.6	0.84	1.26	0.8	0.61	
15	3	2	1.05	0.48	0.34	0.3	0.43	0.31	1.24	1.05	1	1.87	2.28	1.51	1.05	
15	3	3	0.86	0.43	0.3	0.27	0.38	0.28	0.96	0.86	0.83	1.41	2.14	1.18	0.86	
15	3	4	0.81	0.41	0.29	0.26	0.37	0.27	0.89	0.81	0.8	1.24	2	1.08	0.81	
15	3	5	0.79	0.4	0.28	0.25	0.36	0.27	0.87	0.79	0.78	1.17	1.88	1.05	0.79	
15	4	2	0.8	0.38	0.25	0.24	0.35	0.26	0.87	0.81	0.79	1.33	1.89	1.15	0.8	
15	4	4	0.67	0.32	0.22	0.2	0.3	0.23	0.71	0.67	0.67	0.95	1.52	0.89	0.67	
15	4	5	0.66	0.31	0.21	0.2	0.3	0.23	0.69	0.66	0.65	0.93	1.44	0.87	0.66	
15	5	2	0.67	0.31	0.21	0.2	0.3	0.22	0.71	0.67	0.67	1.02	1.57	0.94	0.67	
15	5	3	0.6	0.28	0.19	0.18	0.27	0.21	0.62	0.6	0.6	0.84	1.35	0.8	0.6	
15	5	4	0.58	0.27	0.18	0.17	0.26	0.2	0.59	0.58	0.58	0.8	1.24	0.77	0.58	
15	5	5	0.56	0.26	0.17	0.17	0.25	0.2	0.58	0.56	0.56	0.78	1.19	0.75	0.56	
20	3	3	0.8	0.39	0.27	0.24	0.36	0.27	0.88	0.8	0.8	1.26	2.15	1.1	0.81	
20	3	4	0.77	0.38	0.26	0.24	0.35	0.26	0.84	0.77	0.76	1.16	1.94	1.04	0.77	
20	3	5	0.76	0.37	0.25	0.23	0.34	0.26	0.81	0.76	0.75	1.11	1.81	1.02	0.76	
20	4	2	0.73	0.34	0.23	0.21	0.32	0.24	0.78	0.74	0.73	1.16	1.87	1.04	0.74	
20	4	3	0.66	0.31	0.21	0.19	0.3	0.23	0.69	0.66	0.66	0.95	1.59	0.9	0.66	
20	4	4	0.64	0.3	0.2	0.19	0.29	0.22	0.66	0.64	0.64	0.9	1.45	0.86	0.64	
20	5	2	0.63	0.28	0.18	0.18	0.27	0.21	0.65	0.63	0.63	0.91	1.5	0.86	0.63	
20	5	3	0.57	0.25	0.17	0.16	0.25	0.2	0.58	0.57	0.57	0.79	1.29	0.77	0.57	
20	5	4	0.55	0.24	0.16	0.15	0.24	0.19	0.56	0.55	0.55	0.76	1.19	0.75	0.55	
20	5	5	0.54	0.24	0.16	0.15	0.24	0.19	0.55	0.54	0.54	0.75	1.16	0.73	0.54	
25	3	2	0.86	0.4	0.27	0.25	0.37	0.28	0.95	0.86	0.85	1.53	2.4	1.27	0.86	
25	3	3	0.78	0.37	0.25	0.23	0.34	0.26	0.83	0.78	0.77	1.17	2.09	1.06	0.77	
25	3	4	0.75	0.35	0.24	0.22	0.33	0.25	0.8	0.75	0.75	1.1	1.89	1.02	0.75	
25	3	5	0.74	0.35	0.24	0.22	0.33	0.25	0.78	0.74	0.73	1.07	1.75	1	0.73	
25	4	3	0.64	0.29	0.19	0.18	0.28	0.22	0.66	0.64	0.64	0.91	1.53	0.87	0.64	
25	4	4	0.62	0.28	0.18	0.18	0.27	0.22	0.64	0.62	0.62	0.87	1.4	0.84	0.62	
25	4	5	0.61	0.27	0.18	0.17	0.27	0.21	0.63	0.61	0.61	0.85	1.35	0.83	0.61	
25	5	2	0.6	0.26	0.17	0.16	0.26	0.2	0.61	0.6	0.6	0.85	1.45	0.82	0.6	
25	5	3	0.55	0.24	0.15	0.15	0.24	0.19	0.56	0.55	0.55	0.77	1.24	0.75	0.55	
25	5	4	0.54	0.23	0.15	0.14	0.23	0.19	0.54	0.54	0.54	0.74	1.16	0.73	0.54	
25	5	5	0.53	0.23	0.15	0.14	0.22	0.19	0.53	0.53	0.53	0.73	1.14	0.72	0.53	

Table A.11: Simulation data example for average confidence interval width of P/T

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	4.15	4.22	4.25	4.22	4.19	4.62	4.14	4.15	4.15	4.11	4.07	4.12	4.15	
5	3	3	3.98	4.1	4.14	4.14	4.08	4.57	3.99	3.98	3.98	3.92	3.85	3.93	3.98	
5	3	4	3.89	4.04	4.12	4.1	4.03	4.54	3.87	3.89	3.89	3.82	3.74	3.84	3.89	
5	3	5	3.84	3.99	4.09	4.07	3.99	4.52	3.82	3.84	3.85	3.75	3.67	3.77	3.84	
5	4	2	2.06	2.07	2.08	2.06	2.05	2.36	2.07	2.06	2.06	2.07	2.06	2.06	2.06	
5	4	3	1.97	2.01	2.03	2.02	2	2.33	1.98	1.97	1.97	1.97	1.94	1.96	1.97	
5	4	4	1.93	1.98	2.02	2	1.98	2.32	1.93	1.92	1.92	1.91	1.89	1.91	1.92	
5	5	2	1.38	1.37	1.38	1.37	1.36	1.62	1.39	1.38	1.37	1.4	1.4	1.39	1.38	
5	5	3	1.32	1.33	1.35	1.34	1.32	1.6	1.33	1.32	1.31	1.33	1.33	1.32	1.32	
5	5	4	1.29	1.32	1.33	1.33	1.31	1.6	1.29	1.29	1.29	1.29	1.29	1.28	1.29	
5	5	5	1.27	1.3	1.33	1.32	1.3	1.59	1.27	1.27	1.27	1.27	1.26	1.26	1.27	
10	3	2	3.45	3.33	3.28	3.22	3.27	4.05	3.5	3.45	3.43	3.57	3.58	3.51	3.45	
10	3	3	3.33	3.26	3.23	3.18	3.2	4.02	3.38	3.34	3.32	3.43	3.43	3.37	3.34	
10	3	4	3.28	3.22	3.19	3.15	3.17	4	3.31	3.28	3.27	3.34	3.36	3.31	3.27	
10	3	5	3.24	3.19	3.18	3.14	3.15	3.99	3.27	3.25	3.23	3.29	3.3	3.27	3.24	
10	4	2	1.68	1.57	1.54	1.51	1.55	2.08	1.7	1.68	1.67	1.77	1.81	1.73	1.68	
10	4	4	1.6	1.52	1.49	1.48	1.5	2.06	1.61	1.6	1.59	1.66	1.71	1.64	1.6	
10	4	5	1.58	1.51	1.48	1.47	1.49	2.06	1.59	1.58	1.58	1.64	1.69	1.62	1.58	
10	5	2	1.12	1.02	0.99	0.98	1.01	1.45	1.13	1.12	1.12	1.19	1.24	1.17	1.12	
10	5	3	1.09	1	0.97	0.96	0.99	1.44	1.1	1.09	1.08	1.15	1.21	1.13	1.09	
10	5	4	1.07	0.99	0.96	0.95	0.98	1.43	1.08	1.07	1.07	1.12	1.18	1.11	1.07	
10	5	5	1.06	0.98	0.96	0.95	0.97	1.43	1.07	1.06	1.06	1.11	1.17	1.1	1.06	
15	3	2	3.14	2.91	2.81	2.75	2.86	3.82	3.19	3.14	3.13	3.31	3.37	3.24	3.15	
15	3	3	3.05	2.86	2.77	2.72	2.8	3.81	3.1	3.05	3.05	3.18	3.26	3.14	3.05	
15	3	4	3.01	2.83	2.75	2.7	2.78	3.79	3.05	3.02	3	3.14	3.21	3.1	3.01	
15	3	5	2.99	2.82	2.73	2.69	2.77	3.78	3.02	2.99	2.99	3.1	3.18	3.06	2.99	
15	4	2	1.53	1.36	1.29	1.27	1.34	1.99	1.56	1.53	1.53	1.64	1.72	1.61	1.53	
15	4	4	1.48	1.32	1.26	1.24	1.31	1.97	1.5	1.48	1.47	1.56	1.65	1.55	1.48	
15	4	5	1.46	1.31	1.25	1.24	1.3	1.97	1.48	1.46	1.46	1.55	1.64	1.54	1.46	
15	5	2	1.03	0.88	0.83	0.82	0.87	1.39	1.04	1.03	1.03	1.11	1.19	1.1	1.03	
15	5	3	1.01	0.86	0.81	0.8	0.86	1.39	1.01	1.01	1.01	1.08	1.17	1.07	1.01	
15	5	4	1	0.86	0.81	0.8	0.85	1.38	1	1	0.99	1.07	1.15	1.06	1	
15	5	5	0.99	0.85	0.8	0.8	0.85	1.38	1	0.99	0.99	1.06	1.15	1.05	0.99	
20	3	3	2.91	2.62	2.5	2.45	2.58	3.69	2.95	2.9	2.89	3.05	3.17	3.01	2.9	
20	3	4	2.87	2.6	2.48	2.43	2.56	3.68	2.91	2.87	2.86	3.01	3.13	2.97	2.87	
20	3	5	2.86	2.59	2.47	2.43	2.56	3.67	2.88	2.85	2.84	2.99	3.1	2.95	2.85	
20	4	2	1.46	1.24	1.15	1.14	1.23	1.94	1.48	1.45	1.46	1.57	1.67	1.55	1.46	
20	4	3	1.43	1.22	1.13	1.12	1.21	1.93	1.44	1.43	1.43	1.54	1.64	1.51	1.43	
20	4	4	1.41	1.21	1.13	1.11	1.2	1.93	1.43	1.41	1.41	1.51	1.63	1.5	1.41	
20	5	2	0.99	0.81	0.74	0.73	0.8	1.37	1	0.99	0.99	1.07	1.17	1.06	0.99	
20	5	3	0.97	0.79	0.73	0.72	0.79	1.36	0.98	0.97	0.97	1.05	1.15	1.04	0.97	
20	5	4	0.96	0.78	0.72	0.71	0.78	1.36	0.96	0.96	0.96	1.04	1.14	1.03	0.96	
20	5	5	0.95	0.78	0.72	0.71	0.78	1.36	0.96	0.95	0.96	1.03	1.14	1.03	0.95	
25	3	2	2.85	2.51	2.35	2.3	2.46	3.63	2.9	2.86	2.85	3.06	3.19	2.99	2.85	
25	3	3	2.8	2.47	2.33	2.27	2.44	3.62	2.84	2.8	2.8	2.98	3.12	2.94	2.81	
25	3	4	2.78	2.45	2.3	2.26	2.42	3.6	2.8	2.78	2.77	2.94	3.07	2.9	2.77	
25	3	5	2.76	2.44	2.29	2.25	2.41	3.6	2.79	2.76	2.76	2.91	3.07	2.89	2.76	
25	4	3	1.39	1.15	1.05	1.03	1.14	1.9	1.4	1.39	1.39	1.5	1.62	1.49	1.39	
25	4	4	1.38	1.14	1.05	1.03	1.13	1.9	1.39	1.38	1.38	1.49	1.61	1.47	1.38	
25	4	5	1.37	1.13	1.04	1.02	1.12	1.9	1.38	1.37	1.37	1.48	1.6	1.46	1.37	
25	5	2	0.96	0.76	0.68	0.68	0.76	1.35	0.96	0.96	0.96	1.05	1.16	1.04	0.96	
25	5	3	0.95	0.75	0.67	0.67	0.75	1.35	0.95	0.95	0.94	1.03	1.14	1.03	0.95	
25	5	4	0.94	0.74	0.67	0.66	0.74	1.34	0.94	0.94	0.94	1.02	1.13	1.02	0.94	
25	5	5	0.93	0.74	0.66	0.66	0.74	1.35	0.94	0.93	0.94	1.02	1.13	1.01	0.94	

Table A.12: Simulation data example for average confidence interval width of SNR

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	3.16	3.23	3.26	3.21	3.13	3.22	3.22	3.16	3.12	3.28	3.15	3.14	3.16	
5	3	3	3.16	3.25	3.27	3.22	3.12	3.22	3.19	3.16	3.12	3.2	3.07	3.11	3.16	
5	3	4	3.17	3.25	3.27	3.21	3.12	3.22	3.2	3.17	3.13	3.18	3.04	3.11	3.17	
5	3	5	3.17	3.25	3.27	3.21	3.12	3.21	3.22	3.17	3.13	3.18	3.02	3.12	3.17	
5	4	2	3.05	3.18	3.21	3.11	2.98	3.07	3.13	3.05	2.99	3.15	2.97	3.01	3.05	
5	4	3	3.07	3.18	3.22	3.1	2.96	3.06	3.16	3.07	3	3.12	2.9	3	3.07	
5	4	4	3.07	3.18	3.21	3.1	2.96	3.05	3.18	3.07	2.99	3.15	2.88	3	3.07	
5	5	2	2.98	3.13	3.17	3.02	2.86	2.95	3.11	2.98	2.9	3.09	2.86	2.92	2.98	
5	5	3	2.99	3.12	3.17	3.02	2.85	2.94	3.14	2.99	2.9	3.11	2.8	2.91	2.99	
5	5	4	2.98	3.12	3.16	3.01	2.84	2.93	3.15	2.98	2.89	3.14	2.77	2.91	2.98	
5	5	5	2.97	3.12	3.16	3.01	2.84	2.93	3.15	2.97	2.88	3.15	2.76	2.9	2.97	
10	3	2	2.13	2.21	2.23	2.12	2	2.14	2.2	2.13	2.06	2.19	2.02	2.08	2.12	
10	3	3	2.12	2.2	2.22	2.12	1.98	2.13	2.23	2.12	2.04	2.2	1.97	2.06	2.12	
10	3	4	2.11	2.19	2.21	2.11	1.98	2.12	2.22	2.11	2.03	2.22	1.94	2.05	2.11	
10	3	5	2.1	2.19	2.21	2.11	1.97	2.12	2.22	2.1	2.02	2.22	1.93	2.04	2.1	
10	4	2	1.97	2.09	2.13	1.96	1.78	1.92	2.13	1.97	1.86	2.11	1.82	1.9	1.97	
10	4	4	1.93	2.07	2.11	1.94	1.76	1.9	2.12	1.93	1.82	2.14	1.75	1.86	1.93	
10	4	5	1.92	2.06	2.11	1.94	1.75	1.9	2.11	1.91	1.81	2.14	1.73	1.84	1.91	
10	5	2	1.84	1.99	2.05	1.83	1.63	1.77	2.06	1.84	1.72	2.07	1.69	1.77	1.84	
10	5	3	1.81	1.98	2.04	1.82	1.62	1.76	2.05	1.81	1.69	2.08	1.65	1.74	1.81	
10	5	4	1.79	1.97	2.04	1.81	1.61	1.76	2.03	1.79	1.68	2.07	1.62	1.72	1.79	
10	5	5	1.78	1.97	2.04	1.81	1.61	1.75	2.02	1.78	1.67	2.06	1.61	1.7	1.78	
15	3	2	1.84	1.92	1.94	1.81	1.65	1.83	1.97	1.84	1.75	1.95	1.72	1.79	1.84	
15	3	3	1.82	1.91	1.93	1.8	1.64	1.82	1.97	1.82	1.72	1.98	1.67	1.76	1.81	
15	3	4	1.8	1.9	1.93	1.8	1.63	1.81	1.96	1.8	1.7	1.98	1.64	1.74	1.8	
15	3	5	1.79	1.9	1.93	1.79	1.63	1.8	1.95	1.79	1.69	1.98	1.63	1.72	1.79	
15	4	2	1.64	1.78	1.83	1.61	1.42	1.6	1.87	1.64	1.53	1.88	1.51	1.59	1.65	
15	4	4	1.59	1.76	1.81	1.6	1.4	1.59	1.84	1.59	1.48	1.87	1.45	1.52	1.59	
15	4	5	1.58	1.76	1.81	1.59	1.4	1.58	1.82	1.58	1.47	1.86	1.44	1.51	1.58	
15	5	2	1.5	1.67	1.74	1.47	1.28	1.46	1.77	1.5	1.38	1.82	1.38	1.44	1.5	
15	5	3	1.46	1.65	1.73	1.46	1.27	1.45	1.75	1.46	1.35	1.8	1.35	1.4	1.46	
15	5	4	1.44	1.64	1.73	1.45	1.26	1.45	1.73	1.44	1.34	1.78	1.33	1.38	1.44	
15	5	5	1.43	1.64	1.73	1.45	1.26	1.44	1.72	1.43	1.33	1.76	1.32	1.37	1.43	
20	3	3	1.65	1.76	1.78	1.63	1.45	1.65	1.84	1.65	1.54	1.86	1.52	1.59	1.65	
20	3	4	1.63	1.75	1.78	1.62	1.45	1.65	1.82	1.63	1.53	1.85	1.5	1.57	1.63	
20	3	5	1.62	1.75	1.78	1.62	1.44	1.64	1.81	1.62	1.52	1.85	1.48	1.56	1.62	
20	4	2	1.46	1.62	1.67	1.42	1.24	1.44	1.73	1.46	1.35	1.77	1.35	1.41	1.46	
20	4	3	1.43	1.6	1.67	1.41	1.23	1.44	1.7	1.43	1.32	1.75	1.32	1.37	1.43	
20	4	4	1.41	1.6	1.66	1.41	1.22	1.43	1.69	1.41	1.31	1.73	1.31	1.35	1.41	
20	5	2	1.31	1.49	1.58	1.27	1.11	1.31	1.62	1.31	1.21	1.68	1.23	1.27	1.31	
20	5	3	1.28	1.47	1.57	1.26	1.1	1.3	1.59	1.28	1.19	1.65	1.21	1.24	1.28	
20	5	4	1.26	1.46	1.57	1.25	1.09	1.3	1.57	1.26	1.18	1.62	1.19	1.22	1.26	
20	5	5	1.25	1.46	1.56	1.25	1.09	1.3	1.56	1.25	1.17	1.61	1.18	1.21	1.25	
25	3	2	1.57	1.67	1.7	1.52	1.35	1.56	1.77	1.57	1.46	1.79	1.46	1.52	1.57	
25	3	3	1.54	1.66	1.7	1.51	1.34	1.55	1.75	1.54	1.43	1.79	1.42	1.48	1.54	
25	3	4	1.52	1.66	1.69	1.5	1.33	1.54	1.73	1.52	1.42	1.78	1.4	1.46	1.52	
25	3	5	1.51	1.65	1.69	1.5	1.33	1.54	1.73	1.51	1.41	1.76	1.39	1.45	1.51	
25	4	3	1.31	1.49	1.57	1.28	1.12	1.34	1.6	1.31	1.21	1.66	1.24	1.27	1.31	
25	4	4	1.29	1.48	1.56	1.27	1.11	1.34	1.58	1.29	1.2	1.64	1.22	1.25	1.29	
25	4	5	1.28	1.48	1.56	1.27	1.11	1.34	1.57	1.28	1.2	1.63	1.21	1.24	1.28	
25	5	2	1.19	1.37	1.47	1.13	1	1.22	1.51	1.19	1.11	1.59	1.14	1.17	1.19	
25	5	3	1.17	1.35	1.46	1.12	0.99	1.22	1.47	1.17	1.09	1.55	1.12	1.14	1.17	
25	5	4	1.15	1.34	1.45	1.12	0.99	1.21	1.46	1.15	1.08	1.52	1.11	1.12	1.15	
25	5	5	1.14	1.33	1.45	1.11	0.99	1.21	1.44	1.14	1.07	1.5	1.1	1.12	1.14	

Table A.13: Simulation data example for average confidence interval width of C_p

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	1.06	1.06	1.07	1.07	1.07	1.12	1.06	1.06	1.05	1.07	1.05	1.05	1.05	1.06
5	3	3	1.03	1.05	1.05	1.06	1.06	1.11	1.04	1.03	1.03	1.04	1.02	1.03	1.03	1.03
5	3	4	1.02	1.04	1.05	1.05	1.05	1.11	1.02	1.02	1.02	1.02	1.01	1.02	1.02	1.02
5	3	5	1.02	1.04	1.05	1.05	1.05	1.11	1.02	1.01	1.02	1.01	1	1.01	1.01	1.02
5	4	2	0.88	0.89	0.89	0.9	0.92	0.96	0.88	0.88	0.89	0.88	0.89	0.88	0.88	0.88
5	4	3	0.86	0.87	0.88	0.89	0.92	0.95	0.86	0.86	0.87	0.86	0.86	0.86	0.86	0.86
5	4	4	0.85	0.86	0.87	0.89	0.91	0.95	0.84	0.85	0.86	0.84	0.85	0.85	0.85	0.85
5	5	2	0.79	0.79	0.79	0.81	0.86	0.87	0.77	0.79	0.81	0.77	0.8	0.79	0.79	0.79
5	5	3	0.77	0.77	0.78	0.81	0.85	0.86	0.75	0.77	0.79	0.75	0.78	0.77	0.77	0.77
5	5	4	0.76	0.77	0.77	0.8	0.85	0.86	0.74	0.76	0.79	0.73	0.77	0.76	0.76	0.76
5	5	5	0.75	0.76	0.77	0.8	0.85	0.86	0.73	0.75	0.78	0.72	0.76	0.75	0.75	0.75
10	3	2	0.95	0.94	0.93	0.92	0.91	1.02	0.96	0.95	0.93	0.98	0.96	0.95	0.95	0.95
10	3	3	0.93	0.93	0.92	0.91	0.9	1.01	0.95	0.93	0.92	0.96	0.94	0.93	0.93	0.93
10	3	4	0.92	0.92	0.92	0.91	0.9	1.01	0.94	0.92	0.91	0.95	0.92	0.92	0.92	0.92
10	3	5	0.92	0.92	0.92	0.91	0.9	1.01	0.93	0.92	0.91	0.94	0.91	0.92	0.92	0.92
10	4	2	0.75	0.73	0.72	0.71	0.72	0.83	0.76	0.75	0.74	0.79	0.77	0.75	0.75	0.75
10	4	4	0.72	0.71	0.71	0.7	0.71	0.83	0.74	0.72	0.72	0.76	0.74	0.73	0.72	0.72
10	4	5	0.72	0.71	0.7	0.7	0.7	0.82	0.73	0.72	0.71	0.75	0.73	0.72	0.72	0.72
10	5	2	0.63	0.6	0.59	0.6	0.63	0.72	0.64	0.63	0.63	0.66	0.66	0.64	0.63	0.63
10	5	3	0.62	0.59	0.59	0.59	0.62	0.72	0.63	0.62	0.62	0.65	0.65	0.63	0.62	0.62
10	5	4	0.61	0.59	0.58	0.59	0.62	0.72	0.62	0.61	0.61	0.64	0.64	0.62	0.61	0.61
10	5	5	0.6	0.58	0.58	0.58	0.62	0.72	0.61	0.6	0.61	0.63	0.63	0.62	0.6	0.6
15	3	2	0.9	0.87	0.86	0.85	0.84	0.97	0.92	0.9	0.88	0.94	0.92	0.91	0.9	0.9
15	3	3	0.88	0.87	0.86	0.84	0.83	0.97	0.91	0.88	0.87	0.92	0.9	0.89	0.88	0.88
15	3	4	0.88	0.86	0.85	0.84	0.83	0.97	0.9	0.88	0.86	0.92	0.89	0.88	0.88	0.88
15	3	5	0.87	0.86	0.85	0.84	0.82	0.97	0.89	0.87	0.86	0.91	0.88	0.87	0.87	0.87
15	4	2	0.69	0.66	0.64	0.63	0.63	0.79	0.72	0.69	0.68	0.75	0.73	0.71	0.69	0.69
15	4	4	0.67	0.65	0.63	0.62	0.62	0.78	0.7	0.68	0.66	0.72	0.7	0.69	0.67	0.67
15	4	5	0.67	0.64	0.63	0.62	0.62	0.78	0.69	0.67	0.66	0.72	0.7	0.68	0.67	0.67
15	5	2	0.58	0.53	0.52	0.51	0.54	0.68	0.6	0.58	0.57	0.63	0.62	0.59	0.58	0.58
15	5	3	0.57	0.53	0.51	0.51	0.53	0.68	0.59	0.57	0.56	0.62	0.61	0.58	0.57	0.57
15	5	4	0.56	0.52	0.51	0.5	0.53	0.67	0.58	0.56	0.55	0.61	0.6	0.58	0.56	0.56
15	5	5	0.56	0.52	0.51	0.5	0.53	0.67	0.58	0.56	0.55	0.61	0.6	0.58	0.56	0.56
20	3	3	0.86	0.83	0.81	0.8	0.78	0.94	0.88	0.85	0.83	0.9	0.87	0.86	0.85	0.85
20	3	4	0.85	0.82	0.81	0.79	0.78	0.94	0.87	0.85	0.83	0.89	0.86	0.85	0.85	0.85
20	3	5	0.85	0.82	0.81	0.79	0.78	0.94	0.87	0.84	0.82	0.89	0.86	0.85	0.84	0.84
20	4	2	0.66	0.62	0.6	0.58	0.58	0.76	0.69	0.66	0.65	0.73	0.7	0.68	0.66	0.66
20	4	3	0.65	0.61	0.59	0.57	0.57	0.76	0.68	0.65	0.64	0.72	0.69	0.67	0.65	0.65
20	4	4	0.65	0.61	0.59	0.57	0.57	0.76	0.68	0.65	0.63	0.71	0.69	0.66	0.65	0.65
20	5	2	0.55	0.5	0.47	0.46	0.48	0.65	0.58	0.55	0.54	0.61	0.6	0.57	0.55	0.55
20	5	3	0.54	0.49	0.47	0.46	0.48	0.65	0.57	0.54	0.53	0.6	0.59	0.56	0.54	0.54
20	5	4	0.54	0.48	0.46	0.45	0.48	0.65	0.56	0.54	0.53	0.6	0.58	0.56	0.54	0.54
20	5	5	0.53	0.48	0.46	0.45	0.48	0.65	0.56	0.53	0.52	0.6	0.58	0.55	0.53	0.53
25	3	2	0.84	0.81	0.79	0.77	0.75	0.93	0.87	0.85	0.82	0.9	0.87	0.86	0.84	0.84
25	3	3	0.83	0.8	0.79	0.76	0.75	0.92	0.86	0.83	0.81	0.89	0.86	0.84	0.83	0.83
25	3	4	0.83	0.8	0.78	0.76	0.74	0.92	0.86	0.83	0.81	0.88	0.85	0.84	0.83	0.83
25	3	5	0.82	0.79	0.78	0.76	0.74	0.92	0.85	0.82	0.8	0.88	0.85	0.83	0.83	0.83
25	4	3	0.64	0.59	0.56	0.54	0.54	0.74	0.67	0.64	0.61	0.7	0.68	0.65	0.64	0.64
25	4	4	0.63	0.58	0.56	0.54	0.54	0.74	0.66	0.63	0.61	0.7	0.67	0.65	0.63	0.63
25	4	5	0.63	0.58	0.56	0.54	0.54	0.74	0.66	0.63	0.61	0.7	0.67	0.65	0.63	0.63
25	5	2	0.53	0.47	0.44	0.43	0.45	0.64	0.56	0.53	0.52	0.6	0.58	0.55	0.53	0.53
25	5	3	0.53	0.47	0.44	0.43	0.45	0.64	0.55	0.53	0.51	0.59	0.58	0.55	0.53	0.53
25	5	4	0.52	0.46	0.44	0.42	0.44	0.63	0.55	0.52	0.51	0.59	0.57	0.54	0.52	0.52
25	5	5	0.52	0.46	0.43	0.42	0.44	0.63	0.55	0.52	0.5	0.59	0.57	0.54	0.52	0.52

Table A.14: Simulation data example for average confidence interval width of C_p^*

p	o	n	σ_{Rep}^2	1	1	1	1	1	1	3	3	3	5	5	5	5
			σ_p^2	1	1	1	3	5	5	1	3	5	1	5	5	5
			σ_o^2	1	1	1	1	1	5	3	3	3	5	5	5	5
			σ_{op}^2	1	3	5	5	3	3	3	3	3	3	1	3	5
5	3	2	2	2.43	2.75	2.54	2.01	2.27	2.16	2.03	1.96	2.35	1.92	1.95	2.02	
5	3	3	2.2	2.71	2.98	2.55	2.15	2.3	2.22	2.24	2.1	2.11	1.78	1.99	2.24	
5	3	4	2.42	2.91	2.93	2.65	2.17	2.31	2.52	2.48	2.25	2.18	1.74	2.19	2.41	
5	3	5	2.54	2.78	3.09	2.57	2.09	2.31	2.69	2.47	2.26	2.31	1.76	2.26	2.48	
5	4	2	2.23	2.87	3.14	2.55	2.01	2.14	2.43	2.24	2.09	2.43	1.87	2.02	2.25	
5	4	3	2.47	3.04	3.19	2.6	2.04	2.15	2.82	2.43	2.14	2.5	1.76	2.12	2.41	
5	4	4	2.46	3.1	3.12	2.57	2	2.11	3.04	2.51	2.17	2.76	1.75	2.18	2.51	
5	5	2	2.29	3.01	3.25	2.45	1.93	2.03	2.8	2.28	2.03	2.56	1.83	2.04	2.33	
5	5	3	2.41	2.95	3.21	2.49	1.88	2	3.08	2.43	2.06	2.89	1.71	2.12	2.38	
5	5	4	2.35	2.92	3.1	2.46	1.9	1.99	3.15	2.39	2.05	2.98	1.71	2.1	2.34	
5	5	5	2.37	3.06	3.29	2.48	1.89	1.97	3.17	2.38	2.02	3.16	1.69	2.08	2.35	
10	3	2	2.04	2.74	2.94	2.25	1.53	1.62	2.31	2.02	1.67	1.97	1.36	1.68	2	
10	3	3	2.1	2.73	2.96	2.22	1.53	1.58	2.78	2.1	1.69	2.32	1.28	1.79	2.06	
10	3	4	2.06	2.75	2.99	2.23	1.5	1.55	2.85	2.06	1.64	2.65	1.26	1.75	2.09	
10	3	5	2.01	2.75	2.94	2.25	1.5	1.56	2.88	2.06	1.61	2.9	1.24	1.72	2.06	
10	4	2	1.9	2.68	2.93	2.18	1.34	1.37	2.69	1.88	1.5	2.37	1.25	1.59	1.92	
10	4	4	1.77	2.61	2.87	1.98	1.29	1.32	2.73	1.78	1.39	2.82	1.14	1.49	1.8	
10	4	5	1.76	2.59	2.88	1.97	1.28	1.31	2.68	1.74	1.37	2.79	1.11	1.45	1.75	
10	5	2	1.71	2.47	2.83	1.83	1.21	1.23	2.68	1.71	1.36	2.58	1.18	1.45	1.72	
10	5	3	1.62	2.37	2.8	1.82	1.19	1.21	2.63	1.61	1.29	2.7	1.1	1.4	1.63	
10	5	4	1.58	2.33	2.78	1.8	1.18	1.2	2.52	1.57	1.26	2.68	1.07	1.33	1.57	
10	5	5	1.53	2.41	2.71	1.78	1.18	1.19	2.51	1.54	1.24	2.67	1.05	1.3	1.55	
15	3	2	1.8	2.56	2.82	1.94	1.17	1.2	2.6	1.82	1.37	2.11	1.07	1.44	1.79	
15	3	3	1.72	2.48	2.84	1.92	1.14	1.17	2.68	1.74	1.28	2.64	0.99	1.4	1.71	
15	3	4	1.66	2.5	2.8	1.89	1.13	1.15	2.67	1.66	1.23	2.71	0.95	1.34	1.66	
15	3	5	1.62	2.47	2.77	1.95	1.11	1.13	2.64	1.63	1.21	2.7	0.92	1.28	1.64	
15	4	2	1.49	2.31	2.71	1.65	0.99	1	2.56	1.5	1.11	2.41	0.95	1.23	1.5	
15	4	4	1.34	2.21	2.64	1.57	0.96	0.98	2.41	1.33	1.02	2.51	0.87	1.09	1.34	
15	4	5	1.32	2.27	2.67	1.6	0.96	0.97	2.41	1.31	1.01	2.52	0.85	1.06	1.31	
15	5	2	1.28	2.08	2.49	1.41	0.91	0.91	2.37	1.28	0.99	2.45	0.89	1.08	1.29	
15	5	3	1.2	2.05	2.46	1.39	0.9	0.9	2.25	1.19	0.95	2.4	0.85	1.01	1.2	
15	5	4	1.16	1.98	2.48	1.38	0.89	0.9	2.15	1.15	0.93	2.3	0.83	0.97	1.15	
15	5	5	1.14	1.96	2.45	1.36	0.89	0.89	2.15	1.13	0.92	2.23	0.82	0.95	1.14	
20	3	3	1.42	2.31	2.7	1.69	0.91	0.92	2.49	1.43	1.02	2.61	0.83	1.12	1.43	
20	3	4	1.37	2.32	2.72	1.67	0.9	0.91	2.42	1.36	0.98	2.58	0.78	1.07	1.37	
20	3	5	1.33	2.36	2.7	1.67	0.89	0.91	2.45	1.33	0.97	2.53	0.76	1.02	1.33	
20	4	2	1.21	2.02	2.49	1.36	0.81	0.81	2.31	1.2	0.89	2.43	0.79	0.99	1.2	
20	4	3	1.11	1.98	2.47	1.35	0.8	0.8	2.25	1.11	0.85	2.35	0.75	0.91	1.12	
20	4	4	1.07	1.99	2.47	1.33	0.79	0.79	2.15	1.07	0.83	2.29	0.73	0.87	1.07	
20	5	2	1.02	1.77	2.27	1.14	0.76	0.76	2.1	1.02	0.81	2.29	0.75	0.87	1.02	
20	5	3	0.96	1.72	2.23	1.12	0.75	0.75	1.94	0.96	0.78	2.13	0.72	0.82	0.95	
20	5	4	0.92	1.72	2.2	1.11	0.75	0.76	1.87	0.93	0.77	2.01	0.71	0.8	0.93	
20	5	5	0.91	1.67	2.22	1.1	0.74	0.75	1.82	0.91	0.76	1.98	0.7	0.78	0.91	
25	3	2	1.32	2.17	2.61	1.49	0.79	0.8	2.55	1.31	0.91	2.39	0.76	1.04	1.32	
25	3	3	1.2	2.17	2.59	1.47	0.77	0.78	2.35	1.2	0.85	2.5	0.71	0.93	1.21	
25	3	4	1.15	2.15	2.57	1.45	0.77	0.77	2.25	1.14	0.82	2.41	0.68	0.88	1.15	
25	3	5	1.11	2.11	2.59	1.44	0.76	0.76	2.24	1.12	0.81	2.34	0.67	0.85	1.11	
25	4	3	0.93	1.78	2.32	1.13	0.69	0.69	1.99	0.93	0.73	2.19	0.66	0.77	0.93	
25	4	4	0.89	1.75	2.27	1.12	0.69	0.69	1.93	0.9	0.72	2.14	0.64	0.75	0.9	
25	4	5	0.87	1.75	2.3	1.1	0.69	0.69	1.91	0.88	0.71	2.01	0.63	0.73	0.87	
25	5	2	0.85	1.55	2.09	0.96	0.67	0.67	1.83	0.86	0.7	2.08	0.66	0.75	0.86	
25	5	3	0.81	1.5	2.05	0.94	0.66	0.66	1.7	0.81	0.68	1.92	0.64	0.71	0.81	
25	5	4	0.79	1.48	2.04	0.94	0.66	0.66	1.64	0.78	0.67	1.8	0.62	0.69	0.78	
25	5	5	0.77	1.46	2.03	0.93	0.66	0.66	1.6	0.77	0.67	1.73	0.62	0.68	0.77	

B Off-centering error data

Table B.15: Off-center error for experiment 1

		Replication	
Part	Operator	1	2
A	1	25	20.62
B	1	15	15
C	1	11.18	15
D	1	5	5
E	1	7.07	7.07
F	1	11.18	11.18
G	1	11.18	11.18
H	1	10	5
I	1	5	7.07
J	1	11.18	11.18
A	2	20.62	20
B	2	18.03	15
C	2	20.62	15
D	2	0	5
E	2	7.07	7.07
F	2	5	10
G	2	5	5
H	2	10	10
I	2	5	5
J	2	15.81	15

Table B.16: Off-center error for experiment 2

Part	Operator	Replication	
		1	2
A	1	25	20.62
B	1	15	15
C	1	11.18	15
D	1	5	5
E	1	7.07	7.07
F	1	11.18	11.18
G	1	11.18	11.18
H	1	10	5
I	1	5	7.07
J	1	11.18	11.18
A	2	20.62	20
B	2	18.03	15
C	2	20.62	15
D	2	0	5
E	2	7.07	7.07
F	2	5	10
G	2	5	5
H	2	10	10
I	2	5	5
J	2	15.81	15
A	3	25.5	20.62
B	3	10	15
C	3	15	10
D	3	5	7.07
E	3	7.07	7.07
F	3	5	5
G	3	11.18	11.18
H	3	14.14	10
I	3	5	7.07
J	3	10	7.07

Table B.17: Off-center error for experiment 3

Part	Operator	Replication	
		1	2
A	1	25	20.62
B	1	15	15
C	1	11.18	15
D	1	5	5
E	1	7.07	7.07
F	1	11.18	11.18
G	1	11.18	11.18
H	1	10	5
I	1	5	7.07
J	1	11.18	11.18
A	2	20.62	20
B	2	18.03	15
C	2	20.62	15
D	2	0	5
E	2	7.07	7.07
F	2	5	10
G	2	5	5
H	2	10	10
I	2	5	5
J	2	15.81	15
A	3	25.5	20.62
B	3	10	15
C	3	15	10
D	3	5	7.07
E	3	7.07	7.07
F	3	5	5
G	3	11.18	11.18
H	3	14.14	10
I	3	5	7.07
J	3	10	7.07
A	4	20.62	20
B	4	15	15
C	4	10	15
D	4	5	5
E	4	7.07	7.07
F	4	5	11.18
G	4	5	11.18
H	4	5	5
I	4	5	7.07
J	4	18.03	20.62

C ANOVA tables

Table C.18: ANOVA table for experiment 1

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	P-value
Part	9	1093.13	121.46	9.66	0.00
Operator	1	0.86	0.86	0.07	0.80
Part-Operator	9	113.10	12.57	3.25	0.01
Repeatability	20	77.42	3.87		
Total	39	1284.51			

Table C.19: ANOVA table for experiment 2

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	P-value
Part	9	1485.54	165.06	12.97	0.00
Operator	2	3.67	1.84	0.14	0.87
Part-Operator	18	229.07	12.73	2.90	0.00
Repeatability	30	131.48	4.38		
Total	59	1849.77			

Table C.20: ANOVA table for experiment 3

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	P-value
Part	9	1954.24	217.14	16.10	0.00
Operator	3	3.73	1.24	0.09	0.96
Part-Operator	27	364.11	13.48	2.87	0.00
Repeatability	40	187.86	4.70		
Total	79	2509.94			